

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

CM nucleic - nucleic search, using sw model

Run on: May 22, 2004, 10:57:31 ; Search time 6273 Seconds  
(without alignments)  
10799.484 Million cell updates/sec

Title: US-09-992-901-1

Perfect score: 1563  
Sequence: 1 atggaggaagaagtagcag.....tgaccatcatgaggtatga 1563

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

GenEmbl:  
1: gb\_da:\*  
2: gb\_hg:\*  
3: gb\_in:\*  
4: gb\_cm:\*  
5: gb\_ov:\*  
6: gb\_pat:\*  
7: gb\_ph:\*  
8: gb\_pl:\*  
9: gb\_pr:\*  
10: gb\_ro:\*  
11: gb\_sts:\*  
12: gb\_sy:\*  
13: gb\_un:\*  
14: gb\_vl:\*  
15: em\_da:\*  
16: em\_fun:\*  
17: em\_hum:\*  
18: em\_in:\*  
19: em\_mu:\*  
20: em\_om:\*  
21: em\_or:\*  
22: em\_ov:\*  
23: em\_pat:\*  
24: em\_ph:\*  
25: em\_pl:\*  
26: em\_ro:\*  
27: em\_sts:\*  
28: em\_un:\*  
29: em\_vl:\*  
30: em\_hg\_hum:\*  
31: em\_hg\_inv:\*  
32: em\_hg\_other:\*  
33: em\_hg\_mus:\*  
34: em\_hg\_pin:\*  
35: em\_hg\_rtd:\*  
36: em\_hg\_mam:\*  
37: em\_hg\_vrt:\*  
38: em\_sy:\*  
39: em\_hgo\_hum:\*  
40: em\_hgo\_mus:\*  
41: em\_hgo\_other:\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1559.8	99.8	1563	BT010564	BT010564 Arabidops
2	873.6	55.9	2151	AC003105	AC003105 Arabidops
3	400.6	25.6	2151	AK106672	AK106672 Oryza sat
4	318.8	20.4	1539	AX467615	AX467615 Sequence
5	318.8	20.4	1614	BT003388	BT003388 Arabidops
6	318.8	20.4	1786	AF467303	AF467303 Arabidops
7	291.6	18.7	1548	AF446870	AF446870 Arabidops
8	290.6	18.6	1784	AY052208	AY052208 Arabidops
9	290.2	18.6	1860	AK120018	AK120018 Oryza sat
10	289.6	18.5	1854	CTRCPP450P	CTRCPP450P
11	281.2	18.0	1633	CTRP450B	CTRP450B
12	276.4	17.7	1570	AY066748	AY066748 Arabidops
13	276.4	17.7	1692	AY050827	AY050827 Arabidops
14	273.2	17.5	1572	AB047400	AB047400 Oryza sat
15	273.2	17.5	1903	AK101545	AK101545 Oryza sat
16	272.2	17.4	1813	AB067685	AB067685 Solanum t
17	271.6	17.4	2441	AK105255	AK105255 Oryza sat
18	269.2	17.2	1539	AX467616	AX467616 Sequence
19	269.2	17.2	1539	BT001008	BT001008 Arabidops
20	269.2	17.2	1697	AF367347	AF367347 Arabidops
21	265.8	17.0	1570	BT002745	BT002745 Arabidops
22	265.8	17.0	1618	AY065060	AY065060 Arabidops
23	265.4	17.0	1833	AX467614	AX467614 Sequence
24	259.2	16.6	2227	AF321870	AF321870 Lolium ri
25	257.6	16.5	1666	AF321866	AF321866 Lolium ri
26	257.6	16.5	1739	AF321869	AF321869 Lolium ri
27	253.2	16.2	1821	AF321868	AF321868 Lolium ri
28	253.2	16.2	1865	AF321867	AF321867 Lolium ri
29	250.4	16.0	1620	AY437442	AY437442 Triticum
30	248	15.9	1767	AF325226	AF325226 Nicotiana p
31	246.6	15.8	1963	AK101750	AK101750 Oryza sat
32	246.6	15.8	1811	AK101667	AK101667 Oryza sat
33	245.6	15.7	2098	AK073238	AK073238 Oryza sat
34	244.4	15.6	1587	AB038597	AB038597 Oryza sat
35	241.2	15.4	1557	AX050397	AX050397 Sequence
36	236.4	15.1	1734	AF249329	AF249329 Lycopersi
37	235.8	15.1	1849	AK101513	AK101513 Oryza sat
38	235.8	15.1	2167	AK120700	AK120700 Oryza sat
39	230.6	14.8	1560	AF093168	AF093168 Arabidops
40	230.6	14.8	1855	AF367329	AF367329 Arabidops
41	220.8	14.1	67349	AP003752	AP003752 Oryza sat
42	220.8	14.1	101607	AP004334	AP004334 Oryza sat
43	216.4	13.8	1901	AK074025	AK074025 Oryza sat
44	216.2	13.8	1838	AK072220	AK072220 Oryza sat
45	207.6	13.3	1815	AK107349	AK107349 Oryza sat

ALIGNMENTS

RESULT 1  
BT010564 1563 bp mRNA linear PLN 30-AUG-2003  
LOCUS Arabidopsis thaliana At2g26710 gene, complete cds.  
DEFINITION Arabidopsis thaliana At2g26710 gene, complete cds.  
ACCESSION BT010564.1 GI:34365750  
VERSION BT010564.1  
KEYWORDS FLI CDNA.  
SOURCE Arabidopsis thaliana (thale cress)  
ORGANISM Arabidopsis thaliana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
Rosidae; eustosids II; Brassicales; Brassicaceae; Arabidopsis.  
1 (bases 1 to 1563)  
CHECK, R., Chen, H., Kim, C.J., Shin, P., Bowser, L., Carninci, P.,  
Dale, J.M., Hayashizaki, Y., Ishida, J., Jones, T., Kamiya, A.,

TITLE  
JOURNAL  
AUTHORS  
REFERENCE

Karlín-Neumann, G., Kawai, J., Lam, B., Lin, J., Miranda, M., Narusaka, M., Nguyen, M., Onodera, C.S., Palm, C.J., Quach, H.L., Sakurai, T., Satou, M., Seki, M., Southwick, A., Toriumi, M., Wong, C., Wu, H.C., Yamada, K., Yu, G., Shinozaki, K., Davis, R.W., Theologis, A., and Ecker, J.R.  
Arabidopsis ORF clones  
Unpublished  
2 (bases 1 to 1563)  
Chen, R., Chen, H., Kim, C.J., Shim, P., Bowser, L., Carninci, P., Dale, J.M., Hayashizaki, Y., Ishida, J., Jones, T., Kamiya, A., Karlín-Neumann, G., Kawai, J., Lam, B., Lin, J., Miranda, M., Narusaka, M., Nguyen, M., Onodera, C.S., Palm, C.J., Quach, H.L., Sakurai, T., Satou, M., Seki, M., Southwick, A., Toriumi, M., Wong, C., Wu, H.C., Yamada, K., Yu, G., Shinozaki, K., Davis, R.W., Theologis, A., and Ecker, J.R.  
Direct Submission  
Submitted (30-AUG-2003) Salk Institute Genomic Analysis Laboratory (SIGAL), Plant Biology Laboratory, The Salk Institute for Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037, USA  
The discrepancy does not affect the protein sequence  
RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RAPL cDNAs (RAPL cDNA: 'RIKEN Arabidopsis Full-length cDNA') : Seki, M., Narusaka, M., Ishida, J., Satou, M., Kamiya, A., Sakurai, T., Carninci, P., Kawai, J., Hayashizaki, Y., and Shinozaki, K.  
The Salk, Stanford, PEGC (SSP) Consortium members constructed and sequenced the PUNI (ORF) clones using the RAPL cDNAs: Chen, R., Chen, H., Kim, C.J., Shim, P., Bowser, L., Chan, M.M., Chang, C.W., Dale, J.M., Hsuan, V.W., Jones, T., Karlín-Neumann, G., Lam, B., Lee, J.M., Lin, J., Miranda, M., Nguyen, M., Onodera, C.S., Palm, C.J., Quach, H.L., Southwick, A., Tang, C.C., Toriumi, M., Wong, C., Wu, H.C., Yamada, K., Yu, G., Yuan, S., Davis, R.W., Theologis, A., and Ecker, J.R.  
Chen, R. (SSP/Salk) and Seki, M. (RIKEN GSC) contributed equally to this work. Shinozaki, K. (RIKEN GSC) and Ecker, J.R. (SSP/Salk) contributed equally to this work as PIs.  
Location/Qualifiers  
1..1563  
/organism="Arabidopsis thaliana"  
/mol\_type="mRNA"  
/db\_xref="taxon:3702"  
/chromosome="2"  
/clone="U61967"  
/ecotype="Columbia"  
/note="This clone is in PUNI 51"  
1..1563  
/note="cytochrome P450"  
/codon\_start=1  
/product="At2g26710"  
/protein\_id="AA05187.1"  
/db\_xref="GI:34365751"  
/translation="MEPSSWFIPIPVILVSLIVIKGSLIMBPKIEHFSK  
QGIKGPPIFIGNVKEIVGMKASHPKSHNPVLSFYHMKIKVATLW  
FGTFRITLVAADPILREIFSKSEFESEHNPVYKLEBDDSLSGKGMHRRKILIS  
PTFEMENIKLIVPVILKSVTDVWDKMSDKSENGEVENVIEFQLTEDVSRIRAG  
SSYEDRAVPRILQAOMLCAEAFQKVPFIPGYAFPTFRGNLSKWDKIEIRSLKLI  
ERRONAIIDGEGECKEPALDGLMIQAKNTVODIYEECKSPFADKQPTSMILT  
WTTLLSMHPKMKARDEVILRVGSDVPTKDHVVKLTLSMINSERLTPVAT  
ITRKSIVKIGGYKIPGCTELLIPITAVHDDAIWGDVNEPAPFADGVRAKHP  
VGIPIFGVATCIQNALITLQAKLTIAWIGQFTFLHAPTYQHPATVIMLYPQHA  
PIIFRRLTNHED"  
1356  
/note="compared to genome"  
/replac="C"

ORIGIN

Query Match 99.8%; Score 1559.8; DB 8; Length 1563;  
Best Local Similarity 99.9%; Freq. No. 0;  
Matches 1561; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGGAGAGAAAGATGAGAGAGTGGTTCATTCAGAAAGTTCCTGTCGTATCTTA 60  
DB 1 ATGGAGAGAAAGATGAGAGAGTGGTTCATTCAGAAAGTTCCTGTCGTATCTTA 60  
QY 61 AGCTCTGTAATGAGAAAGGATATCTCTGTTATGTTGAGACCAAGAAAGATTGAAGA 120  
DB 61 AGCTCTGTAATGAGAAAGGATATCTCTGTTATGTTGAGACCAAGAAAGATTGAAGA 120  
QY 121 CATTCCTGTAACAAAGAGATTCAGAGTCCCTTATCATTCATTCATTCGAAATGTAA 180  
DB 121 CATTCCTGTAACAAAGAGATTCAGAGTCCCTTATCATTCATTCATTCGAAATGTAA 180  
QY 181 GAACCTGTTGGAATGATGCTTAAAGCTTCTTCATCTATGCTTCTCTCACATATT 240  
DB 181 GAACCTGTTGGAATGATGCTTAAAGCTTCTTCATCTATGCTTCTCTCACATATT 240  
QY 241 CTTCCTAGAGTTCCTCTTTTACATCATCGAGAAATCTACAGGTGCTACATTTCTG 300  
DB 241 CTTCCTAGAGTTCCTCTTTTACATCATCGAGAAATCTACAGGTGCTACATTTCTG 300  
QY 301 GTTGGTTCGATCCAACTTTCGCTTAAAGGTACCGATCTGATTTGATCAGAGATC 360  
DB 301 GTTGGTTCGATCCAACTTTCGCTTAAAGGTACCGATCTGATTTGATCAGAGATC 360  
QY 361 TTCTCTAAGTCTGATTTCTAGAGAAAGATGAGTCAACCTTGGTTAAACAATTGAA 420  
DB 361 TTCTCTAAGTCTGATTTCTAGAGAAAGATGAGTCAACCTTGGTTAAACAATTGAA 420  
QY 421 GGGAGATGAGCTACTAGTCTCAAGGTGAAATAGGGCTCATCTCAAGAAATCATTAAGC 480  
DB 421 GGGAGATGAGCTACTAGTCTCAAGGTGAAATAGGGCTCATCTCAAGAAATCATTAAGC 480  
QY 481 CCTACTTTCATATGAGAAATCTTAAGTCTTGTACCAAGTGTGTTGAAGAGTGAAT 540  
DB 481 CCTACTTTCATATGAGAAATCTTAAGTCTTGTACCAAGTGTGTTGAAGAGTGAAT 540  
QY 541 GATATGATGAGTAAATAGGTCGATAGTATATCAAGAAACGCTGAAATGATGATGTC 600  
DB 541 GATATGATGAGTAAATAGGTCGATAGTATATCAAGAAACGCTGAAATGATGATGTC 600  
QY 601 TATGATGATGTTCAAGTTTGAATGATGATGATGATGATGATGATGATGATGATGATG 660  
DB 601 TATGATGATGTTCAAGTTTGAATGATGATGATGATGATGATGATGATGATGATGATG 660  
QY 661 TATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 720  
DB 661 TATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 720  
QY 721 GCTTTTCAAAAAGTCTTCAATTCCTGCTATGATTTTTCGCAAGAGGAAATTGAG 780  
DB 721 GCTTTTCAAAAAGTCTTCAATTCCTGCTATGATTTTTCGCAAGAGGAAATTGAG 780  
QY 781 TCTCGAAGTGTAGCAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATG 840  
DB 781 TCTCGAAGTGTAGCAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATG 840  
QY 841 CAAGAAGTATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 900  
DB 841 CAAGAAGTATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 900  
QY 901 GATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 960  
DB 901 GATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 960  
QY 961 TTTTCTTCCGCGGAAAGAGCAACTTCTGATGATGATGATGATGATGATGATGATGATG 1020  
DB 961 TTTTCTTCCGCGGAAAGAGCAACTTCTGATGATGATGATGATGATGATGATGATGATG 1020  
QY 1021 TCCATGACCCCGAGAGTGGCAGGCAAGACCGATGAGAGTCTGAGGCTGAGGCTCA 1080  
DB 1021 TCCATGACCCCGAGAGTGGCAGGCAAGACCGATGAGAGTCTGAGGCTGAGGCTCA 1080  
QY 1081 CGTATGTCCTTACCAAGACATGTCGTTAAGCTTAAACGTTGATGATGATCTTGAAC 1140







Db	52416	GCTATAGAGCGGCGGAGACAAACGCTATAGATGAGGAAGGGAAGATGTAAAGAGCC	52475
Qy	882	GGCGGCGAGAGATTTGTTGGATTAATGATTGAGGCAAGAAATGTGACGGTTCAAGACAT	941
Db	52476	GGCGGCGAGAGATTTGTTGGATTAATGATTGAGGCAAGAAATGTGACGGTTCAAGACAT	52535
Qy	942	TGTGAGAGAGTGTAAAGCTTTCTTCGCGGGGAAACAGACACTTCTATCTGTGAC	1001
Db	52536	TGTGAGAGAGTGTAAAGCTTTCTTCGCGGGGAAACAGACACTTCTATCTGTGAC	52595
Qy	1002	GTGAGAGACATCTTGTCTATCATGACACCGGAGTGGAGCCAAAGCAGTGAAGGT	1061
Db	52596	GTGAGAGACATCTTGTCTATCATGACACCGGAGTGGAGCCAAAGCAGTGAAGGT	52655
Qy	1062	CCTGAGGCTTCGGCGCTCAGTGATGTCCCTACCAAGGACCATGTGGTTAAGCTTAAAC	1121
Db	52656	CCTGAGGCTTCGGCGCTCAGTGATGTCCCTACCAAGGACCATGTGGTTAAGCTTAAAC	52715
Qy	1122	-----	1121
Db	52716	GGTGGCTTCTCTTCTGCTTCTGCTTTGTTTCTTCAATTTGGAAATATATCATTA	52775
Qy	1122	-----	1121
Db	52776	ATTACAGATAAAAGGCGAATTTGTGATAGACAGTTTCGTTACCGTTAATTATGATCTT	52835
Qy	1122	-----	1121
Db	52836	TTGTGCGCTAATCTATTAATTTTATATGCGCTATATCTACATCCATGCAAGCAAGTT	52895
Qy	1122	-----	1121
Db	52896	TCAAGGCAAAAGGATATATGATAGACAGATATGTTGATGTTTGTACCTCTTTCA	52955
Qy	1122	-----	1121
Db	52956	AAATTTGAGTGTAGATATATCTTGAACAGCTCTTTAAGTTGTTATCCACCAATAGTAGCT	53015
Qy	1174	ACGATTGACAGCGGCTAAATCCGATGTGAAGCTAGAGAGGTACAAATCCATGTGACG	1233
Db	53016	ACGATTGACAGCGGCTAAATCCGATGTGAAGCTAGAGAGGTACAAATCCATGTGACG	53075
Qy	1234	GAGCTTTCAATCCCATATATGACCGGTCTATATGACCAAGCATTGGGGTATATGACGT	1293
Db	53076	GAGCTTTCAATCCCATATATGACCGGTCTATATGACCAAGCATTGGGGTATATGACGT	53135
Qy	1294	AACGAATTCATCCAGCTCGGTTTGGAGTGAAGTGCAGCTGCTGCAACACCCCGTT	1353
Db	53136	AACGAATTCATCCAGCTCGGTTTGGAGTGAAGTGCAGCTGCTGCAACACCCCGTT	53195
Qy	1354	GCGCTTCATACCGTTTGGCTTGGAGTGTATCATGATGTGTGCAATCTTGTCTATCTT	1413
Db	53196	GCGCTTCATACCGTTTGGCTTGGAGTGTATCATGATGTGTGCAATCTTGTCTATCTT	53255
Qy	1414	CAGGCGCAATTCAGACCTCGCTGTATGATTCACACGCTTCACTTCACTTGGCTCTACT	1473
Db	53256	CAGGCGCAATTCAGACCTCGCTGTATGATTCACACGCTTCACTTCACTTGGCTCTACT	53315
Qy	1474	TATCAGCATGACACCTACCGTCTTATGTTGCTTATCTCTCAACATGTGTGCAACATACC	1533
Db	53316	TATCAGCATGACACCTACCGTCTTATGTTGCTTATCTCTCAACATGTGTGCAACATACC	53375
Qy	1534	TTCCGAGATTTAGACCAATCATGAGATTGA	1563
Db	53376	TTCCGAGATTTAGACCAATCATGAGATTGA	53405

RESULT 3  
AKI06672  
LOCUS AKI06672 2151 bp mRNA linear PLN 24-JUL-2003  
DEFINITION *Oryza sativa* (japonica cultivar-group) cDNA clone:002-114-A02, full  
insert sequence.

AKI06672	GI:32991861
AKI06672.1	Full cDNA; oligo capping.
KEYWORDS	<i>Oryza sativa</i> (japonica cultivar-group)
SOURCE	<i>Oryza sativa</i> (japonica cultivar-group)
ORGANISM	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehharctoideae; Oryzaceae; <i>Oryza</i> .
REFERENCE	1
AUTHORS	The Rice Full-length cDNA Consortium, National Institute of Agrobiological Sciences Rice Full-length cDNA Project Team; Kikuchi, S., Satoh, K., Nagata, T., Kawasashira, N., Doi, K., Kishimoto, N., Yazaki, J., Ishikawa, M., Yamada, H., Ooka, H., Hotta, I., Kojima, K., Namiki, T., Ohneda, E., Yahagi, W., Suzuki, K., Li, C., Ohtsuki, K., Shishiki, T., Foundation of Advancement of International Science Genome Sequencing & Analysis Group; Ohtomo, Y., Iida, Y., Sugano, S., Fujimura, T., Masuda, H., Kobayashi, M., Xie, Q., Lu, M., Kurosaki, T., Kodama, T., Masuda, H., Kobayashi, M., Xie, Q., Lu, M., Narikawa, R., Sugiyama, A., Mizuno, K., Yokomizo, S., Nikura, J., Ikeda, R., Ishibiki, J., Kawamata, M., Matsubara, K., RIKEN, Kusumegi, T., Oka, M., Ryu, R., Ueda, M., Matsubara, K., RIKEN, Kawai, J., Carninci, P., Adachi, J., Aizawa, K., Arakawa, T., Fukuda, S., Hara, A., Hashidume, M., Hayatsu, N., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kondo, S., Kono, H., Miyazaki, A., Otsu, N., Ota, Y., Saito, R., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T., Yoshino, M., and Hayashizaki, Y.
TITLE	Collection, mapping, and annotation of over 28,000 cDNA clones from japonica rice
JOURNAL	Science 301 (5631), 376-379 (2003)
MEDLINE	22752273
PUBMED	12869764
REFERENCE	2 (bases 1 to 2151)
AUTHORS	Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Doi, K., Fujimura, T., Fukuda, S., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayashizaki, Y., Hayatsu, N., Hiraoka, K., Hiraoka, T., Hori, F., Hotta, I., Iida, J., Iida, Y., Ikeda, R., Imamura, K., Imotani, K., Ishibiki, J., Ishii, Y., Ishikawa, M., Itoh, M., Kagawa, I., Kanagawa, S., Katoh, H., Kawasashira, N., Kawai, J., Kawamata, M., Kikuchi, S., Kishikawa-Hirozane, T., Kishimoto, N., Kobayashi, M., Kodama, T., Kojima, K., Kojima, Y., Kondo, S., Kono, H., Koude, M., Koya, S., Kurihara, C., Kurosaki, T., Kusumegi, T., Li, C., Lu, M., Masuda, H., Matsubara, K., Matsuyama, T., Matsuura, Y., Miyazaki, A., Mizuno, K., Murakami, K., Murakami, M., Nagata, T., Nakamura, M., Namiki, T., Narikawa, R., Nikura, J., Nishikawa, K., Nomura, K., Numasaki, R., Ohneda, E., Ohno, M., Ohtsuki, K., Oka, M., Ooka, H., Otsu, N., Ota, Y., Ohtomo, Y., Ryu, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Sato, K., Satoh, K., Shibata, K., Shinagawa, A., Shiraki, T., Shishiki, T., Sogabe, Y., Sugano, S., Sugiyama, A., Suzuki, K., Suzuki, Y., Tagami, M., Tagami-Takeda, Y., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toya, T., Tsunoda, Y., Ueda, M., Waki, K., Xie, Q., Yahagi, W., Yamada, H., Yamamoto, M., Yasunishi, A., Yazaki, J., Yokomizo, S., and Yoshimura, A.
TITLE	Direct Submission
JOURNAL	Submitted (27-AUG-2002) Shoshi Kikuchi, National Institute of Agrobiological Sciences, Department of Molecular Genetics, Head of Laboratory of Gene Expression; 2-1-2 Kannondai, Tsukuba, Ibaraki 305-8602, Japan (E-mail:skikuchi@nias.affrc.go.jp, Tel:81-29-838-7007, Fax:81-29-838-7007)
COMMENT	This clone is one of the 28k full-length cDNA clones from japonica rice. URL: <a href="http://cdna01.dna.affrc.go.jp/cDNA/">http://cdna01.dna.affrc.go.jp/cDNA/</a> NIA Rice Full-length cDNA Project Team: Kikuchi, S., Satoh, K., Nagata, T., Kawasashira, N., Doi, K., Kishimoto, N., Yazaki, J., Ishikawa, M., Yamada, H., Ooka, H., Hotta, I., Kojima, K., Namiki, T., Ohneda, E., Yahagi, W., Suzuki, K., Li, C., Ohtsuki, K., Shishiki, T., and Yamamoto, M.
FAIS	Genome Sequencing & Analysis Group: Ohtomo, Y., Iida, Y., Fujimura, T., Ikeda, R., Ishibiki, J., Kawamata, M., Kobayashi, M., Kodama, T., Kurosaki, T., Kusumegi, T., Lu, M., Masuda, H., Mura, J., Mizuno, K., Narikawa, R., Nikura, J., Oka, M., Ryu, R., Sugano, S., Sugiyama, A., Suzuki, Y., Tsunoda, Y., Ueda, M., Xie, Q., Yokomizo, S., Yoshimura, A., Matsubara, K., and Murakami, K.

Genome Exploration Research Group in Riken Genomic Sciences Center  
and Genome Science Laboratory in Riken: Adachi, J., Aizawa, K.,  
Akumura, T., Arakawa, T., Carninci, P., Fukuda, S., Hara, A.,  
Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K.,  
Hirooka, T., Hori, F., Iida, J., Imamura, K., Imoto, K., Ishii, Y.,  
Itoh, M., Kaga, I., Kanagawa, S., Kato, H., Kawai, J.,  
Kishikawa-Hirozane, T., Kojima, Y., Kondo, S., Kono, H., Kouda, M.,  
Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,  
Nakamura, M., Nishi, K., Nomura, K., Numasaki, R., Ohno, M., Ose, N.,  
Ota, Y., Satoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H.,  
Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T.,  
Sogabe, Y., Tagami, M., Tagami, T., Tanaka, T., Tomaru, A., Toyata, T.,  
Waki, K., Yasunishi, A., and Hayashizaki, Y.

## FEATURES

source

Location/Qualifiers  
1..2151  
/organism="Oryza sativa (japonica cultivar-group)"  
/mol\_type="mRNA"  
/cultiyar="Nidponbare"  
/db\_xref="taxon:39947"  
/clone="002-114-A02"

## ORIGIN

Query Match

25.6%; Score 400.6; DB 8; Length 2151;

Best Local Similarity 58.0%; Pred. No. 1e-103;

Matches 861; Conservative 0; Mismatches 569; Indels 69; Gaps 7;

90 GTTATGCTGGAGACCAAGAAAGATTGAAGACCTTTCTTAACAAGAAATTCAGAGTCC 149  
175 GGTGCTGCTGGAGACCAAGAAAGATTGAAGACCTTTCTTAACAAGAAATTCAGAGTCC 234  
150 TCTTATCATTTCTTCAATCGGAATGTTAAGAATCTTGTGATGATGCTTAAAGCTTC 209  
235 GCCCTACCGCTTCTCGTGGGTGCTGAGAGAGATGAGGCGCTATGCGGAGGACAC 294  
210 TTCTATCTTATGCTTCT 263  
295 CGGAGAGCCATGCG 354  
264 CCATCATCTGAGAGAAATCTACGCTGCTACATTTCTGTTGGTGGTCCAACTTTCG 323  
355 CCATCATCTGAGAGAAATCTACGCTGCTACATTTCTGTTGGTGGTCCAACTTTCG 414  
324 GTTAAAGCTTACCGATCTGATTTGATCAAGAGATCTTCTTAAG---TCTGAGTTCTA 380  
415 GCTCAAGCTGCGAGACCGGAGATGCTGCGAGATCTTCTCAAGCGCGCGCGAGGCGTT 474  
381 CGAGAGAGATGAGCTACCTCTTGTGTTAAACAATTTGAAGCGATGAGCTACTTACT 440  
475 CGAGAGAGATGAGCTACCTCTTGTGTTAAACAATTTGAAGCGATGAGCTACTTACT 534  
441 CAAGAGTGAAGAAATGAGCTGATCAATGAGAAATCAATTAAGCTTCTTCAATGAGAA 500  
535 CCAAGGCGACAAAGGAGCTCAACACGCGCGCTCTCAACCGCGGCTTCTTCCCGACAA 594  
501 TCTTAAGTTCTGTTACAGTTGTTGTTGAAGATGATGATGATGATGATGATGATGAT 560  
595 CCTCAACCGCTGCTGCGCAGCGCGCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 654  
561 CGAT---AGTTATCAAGAAACGTTGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 617  
655 CGCATGCGCTGCG 714  
618 TTGACCTGAAGATGTTATGATGAACAGCTTTTGAAGTGAAGTGAAGTGAAGTGAAG 677  
715 GGTGCGGAGAGGAGCTATCAACGCGCGCAGCTTCCGCGCAGCTACGACTCCGCGCG 774  
678 AGTTTTCGACTTAAGCTCAACAAAGCTTCTTGTGTTGAAGCTTTTCAAAAGTCTT 737  
775 CGTGTTCGCTTGAAGGCGCGCTCATGAGGCTTCCGCGCGCTTCCGCGAGGCTCT 834  
738 CATCTCGCTATGATTTTCTCGACAAAGAGGAAATTTGAAGTCTCGAAAGTTAGACAA 797

835 CGTCCCGGAGTACAGATTCTCTGCGACCAAGAAAGACAGGATGTCGCGGCGCTGACAG 894  
798 GGAAGTAAAGAGCTGTTGTTGAAGCTGATGAGCGCGCGGACAAAGAGCTATGATG 857  
895 GGAAGTAAAGAGCTGTTGTTGAAGCTGATGAGCGCGCGGACAAAGAGCTATGATG 954  
858 AGAAGAGG-----AAGATGTAAGAGCTGCGCGCGGACAAAGATTTGT 899  
955 GGAAGAGAGAGACCAACAGAGCTCAAAAGACAAAGACAGCGGCTTCAACGACTTCT 1014  
900 GGAATTAATGATTA-----GCAAGAAATGACCGCTTCAAGACATTTGAGAG 950  
1015 GGGGCTCATGATCAATGCGCGCGCTGACAGAGACGATGCGGTGAGAGACATGTTGAG 1074  
951 GTGTAAAGCTTTTCTTCCGCGGAAACAGACAACTTCTTATCTGCTGACGTGACGAC 1010  
1075 GTGCAAGACCTTCTTCTTCTGCGCGGACAGACAGACCAACCTGCTCACTGCGGAC 1134  
1011 CATCTGCTATCAATGAGCGCGGAGTGGCGGACCAAGACGATGAGTCTGCGAGGT 1070  
1135 CGTGTGCTGCGCATGACACCGGCGCTGACAGAGCGCGCGCGCGCGCGCGCGCGCG 1194  
1071 CTGCGG-----CTCACGTGATGTCCTTACCAAGACCATGCTTAAAGCTTAAACGTT 1124  
1195 CTGCGGCGATGCGCGCGCGCGGAGTCCCGACCAAGAGACCACTCCCAAGCTCAAGCGCT 1254  
1125 GAGTATGATCTTGAAGACAGCTTTAAGTTGATCAACCAATGATGATGATGATGATG 1184  
1255 CGGATGATCTTCAACGAGACGCTGCGCTGATACCGCGCGGCGGCGGCGGCGGCGG 1314  
1185 CGCTAATCGATGATGAGCTAG-----GAGGTTACAAAT 1220  
1315 CGCAAGTTCAACGCTACCTCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1374  
1221 CGCATGAGCAAGAGCTTCAATCCCAATCAAGAGTGCATGATGATGATGATGATGATG 1280  
1375 CCGCGCGACAGAGAGCTGCTGCTGCGATGATGAGGATGATGATGATGATGATGATG 1434  
1281 GGGTATGAGCGTGAAGCAATTTCAATCGAGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1340  
1435 GGGGCGCGACCG 1494  
1341 CAAGACCGCGTGGCTTCAATGAGCTTGGCTGCGAGTTCGATGATGATGATGATGATG 1400  
1495 GAGAGACCGCGTGGCTTCAATGAGCTTGGCTGCGAGTTCGATGATGATGATGATGATG 1554  
1401 TCTTCTATGATGATGAGCGCAATTTGACATCGCTGATGATGATGATGATGATGATG 1460  
1555 CTTGCGCTCAATCTCAAGAGCTCAACAGAGCGCTTCTTCTCAAGCGCTTCAAGCGCT 1614  
1461 CTGCGCTCAATCTCAAGAGCTCAACAGAGCGCTTCTTCTCAAGCGCTTCAAGCGCT 1520  
1615 GCTCTGCGCGACCTTCAAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1674  
1521 TGACCAATCACTTCCG 1539  
1675 CGCGCGCTGATGATTTCCG 1693

RESULT 4  
AX467615  
LOCUS AX467615 1539 bp DNA linear PAT 16-UTL-2002  
DEFINITION Sequence 3 from Patent WO246377.  
ACCESSION AX467615  
VERSION AX467615.1 GI:21900803  
KEYWORDS  
SOURCE  
ORGANISM Arabidopsis thaliana (thale cress)  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophytes; Magnoliophyta; eudicotyledons; core eudicots;  
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.  
REFERENCE  
AUTHORS Kirochko, J.E., Cutler, A.J. and Abrams, S.R.

TITLE Abscisic acid 8'- and 7'- hydroxylase genes and related sequences  
 from plants  
 JOURNAL Patent: WO 0246377-A 3 13-JUN-2002;  
 National Research Council of Canada (CN)  
 FEATURES Location/Qualifiers  
 source 1.1539  
 /organism="Arabidopsis thaliana"  
 /mol\_type="unassigned DNA"  
 /db\_xref="taxon:3702"  
 /note="var. Columbia"  
 ORIGIN  
 Query Match 20.4%; Score 318.8; DB 6; Length 1539;  
 Best Local Similarity 52.8%; Pred. No. 3,7e-80;  
 Matches 761; Conservative 0; Mismatches 672; Indels 9; Gaps 3;  
 QY 90 GTTATGCTGAGACCAAGAAAGATTGAAACATTTCTTAACAGAGAAATTCAGAGTCC 149  
 DB 78 GGTCTGATTAACCAAGAGTCTTGAGATCTTTGAAAGACAGGGTCTTACCGGAC 137  
 QY 150 TCCCTATCATTTCTTCATCGAAATGTTAAAGACTTTTGGAATGATGCTTAAAGCTTC 209  
 DB 138 TCCCTAACCTCCTCTGCTGAGATATAAGAGAAATGTTATGATGATGAGAGAG 197  
 QY 210 TTCTCATCTATGCTTCTCTCACATATTTCTCTAGAGTTCTCTTTTACCATCA 269  
 DB 198 ATCTAAACCATTAATGTAACGATGATATACCCAGCTCTCTCTCTGCTTAA 257  
 QY 270 CTGAGAGAAATCTACGGTGTACATTTCTGTTTGTGGTCCGTCACATTTCCGTTAC 329  
 DB 258 GATGCTCATTTCTACGGAAGACATTTCTTCAATGATGAGACCATCTCCACAGATGT 317  
 QY 330 GGTAGCCCATCTGATTTATGAGAGATCTTCTCAAGTGTGATTTCTACAGAGAA 389  
 DB 318 GATTAAGAAATCTGACAGATCAAGAGATCTTAAATGAATGATGATGAGAGAA 377  
 QY 390 TGAAGCTCAACCTTTGTTAAACAACTTGAGGCGATGACTTAAGTCTCAAGGTGA 449  
 DB 378 TTCTCATTTCTTGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 434  
 QY 450 AAAATGGCTCATCATGAAAAATCTATGCCCCCTTCTTCAATGAGAAATCTTAAGT 509  
 DB 435 TAAATGGCGAGTCACAGAGAGATCAACCCGCTTTTCACTCGAAAAATCAAGAA 494  
 QY 510 GCTTGAACAGTGTGTGAAGGTGATGATGATGATGATGATGATGATGATGATGATGAT 569  
 DB 495 CATGATCCCTGCTTCTACCATTTGTCAGCGAGGTGTCTGTCATGAGAGAAAGCTATT 554  
 QY 570 ATCAGAAACG--GTGAAGTTAGGTAGATGTTCTATGAGTGTTCAGATTTTGACTGA 626  
 DB 555 TACAGATTAAGAAATGCGCTCTGTAAGTCAATGTTGGCTTGTGAATATATACATGC 614  
 QY 627 AGATGTTATTAAGTAAACAGCTTTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 686  
 DB 615 GGAATGATCTCATCTGCTTTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 674  
 QY 687 ACTCAAGCTCAACAAATGCTTTTGTGCTGAAGCTTTTCAAAAAGTCTTCAATCTCTGG 746  
 DB 675 ACTCAAGGGAATGCTGCTGATGCTTATGCAACAGCTTTTAAAGAAATCTTATACCTCTGG 734  
 QY 747 CTATAGATTTTTCGCAAGAGAGAAATTTGAAGTCTCGAAGTTAGACAAAGAGATAG 806  
 DB 735 ATCAGAGTTTTCACCAAGAGAGATAGAGAGATAGAGAGATAGAGAGATAGAGAGATAG 794  
 QY 807 GAATGCTGTTGAAGTGTATGAGAGCGGAGAGACAAACGCTTATGATGAGAGAGAGAG 866  
 DB 795 CGTAAATTAAGAGATTTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 854  
 QY 867 AGAATGTAAG 923  
 DB 855 AAACGATGATTTGTTGGAGATATGCTTGAATCCATTCAGAGAGAGATTCAGAGAGAGAG 914  
 QY 924 TGTACGGTTCAGAGACATGTTGAGAGAGTAAAGAGCTTTTCTTCCGCGAGAGAAACAGAC 983

DB 915 AATGAGCTGAGAGATGATGAAAGAGTGAAGTGTATTTATTTGCGGAGCAAGAGAC 974  
 QY 984 AACTTTATATCTGCTGAGAGAGAGACCATCTTCTATTCATACACCCGAGTGGAGGC 1043  
 DB 975 AACTTCAGTACTTTTGGTGTGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 1034  
 QY 1044 CAAGACAGTATGAGAGTCCGAGGAGTCCGAGTCAAGTATGATGATGATGATGATGATGATGATGATGATGATGAT 1103  
 DB 1035 TCGGACAGAGAGAGAGAT 1094  
 QY 1104 TGTGTTAAGCTTAAACGTTGAT 1163  
 DB 1095 CTTTACACACTTAAAGTCAATGAT 1154  
 QY 1164 AATGAT 1223  
 DB 1155 GGTGCTCAGCTTAAAGAGTGTCAACAAAGAAATGAT 1214  
 QY 1224 ATGTGACAGAGCTTCTAATCCCATATAGCGGTCCATCATGACCAAGCATTTGGGG 1283  
 DB 1215 AGCTGATTTCAATTAAT 1274  
 QY 1284 TAATGAT 1343  
 DB 1275 CGAT 1334  
 QY 1344 AATCCCGTGGCTTATACCGTGTGGCTGAGAT 1403  
 DB 1335 GAACAGATCTCTTCTTCCCTTTGATGAGGAGCCTAAGATTTGAT 1394  
 QY 1404 TGTATATCTTCAAGGCAATTAATGAT 1463  
 DB 1395 TGTATATGTTGAGGCAATTAATGAT 1454  
 QY 1464 GGTCTCTCTTATGAT 1523  
 DB 1455 CTCTCTCTTATGAT 1514  
 QY 1524 AC 1525  
 DB 1515 TC 1516  
 RESULT 5  
 BT003388 1614 bp mRNA linear PLN 29-JAN-2003  
 LOCUS Arabidopsis thaliana cytochrome P450 (At3g14610) mRNA, complete cds.  
 DEFINITION  
 ACCESSION BT003388  
 VERSION BT003388.1 GI:28059361  
 KEYWORDS Full cDNA.  
 SOURCE Arabidopsis thaliana (thale cress)  
 ORGANISM Arabidopsis thaliana  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eucosids II; Brassicales; Brassicaceae; Arabidopsis.  
 1 (bases 1 to 1614)  
 Soutwicz, A., Nguyen, M., Tripp, M., Palm, C. J., Jones, T., Wu, T., Carninci, P., Chen, H., Cheuk, R., Chan, M. M., Chang, C. H., Dale, J. M., Deng, J. M., Hayashizaki, Y., Hsuan, V. W., Lee, J. M., Ishida, J., Kamiya, A., Kawai, J., Kim, C. J., Natusaka, M., Onodera, C. S., Quach, H. L., Sakurai, T., Saitou, M., Seki, M., Shim, P., Tang, C. C., Toriumi, M., Wong, C., Wu, H. C., Yamada, K., Yu, G., Yuan, S., Shinozaki, K., Becker, J., Theologis, A. and Davis, R. W.  
 Direct Submission  
 Submitted (29-JAN-2003) DNA Sequencing and Technology Center, Stanford University, 855 California Avenue, Palo Alto, CA 94304, USA  
 e-mail for correspondence: arab@sequence.stanford.edu  
 COMMENT  
 RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RFLV cDNAs (RFLV cDNA: "RIKEN

Arabidopsis Full-length cDNA: Seki, M., Narusaka, M., Ishida, J., Satou, M., Kamliya, A., Sakurai, T., Carninci, P., Kawai, J., Hayashizaki, Y., and Shinzaki, K.

The Salik, Stanford, PGSC (SSP) Consortium members carried out the sequencing and annotation of the RAPL cDNAs: Nguyen, M., Tripp, M., Southwick, A., Palm, C. J., Jones, T., Wu, T., Chen, H., Cheuk, R., Chan, M. M., Chang, C. H., Dale, J. M., Deng, J. M., Heuan, V. N., Lee, J. M., Kim, C. J., Quach, H. L., Odoera, C. S., Shim, P., Tang, C. C., Tourant, M., Wong, C., Wu, H. C., Yamada, K., Yu, G., Yuan, S., Ecker, J., Theologis, A., and Davis, R. W.

Nguyen, M., (SSP/Stanford) and Seki, M. (RIKEN GSC) contributed equally to this work. Shinzaki, K. (RIKEN GSC) and Davis, R. W. (SSP/Stanford) contributed equally to this work as PI.

## FEATURES

## Source

Location/Qualifiers

1..1614

/organism="Arabidopsis thaliana"

/mol\_type="mRNA"

/db\_xref="taxon:3702"

/chromosome="3"

/clone="U15533"

/ecotype="Columbia"

/note="This clone is in pENTR/SD-topo. This is a cloned PCR product using RIKEN clone RAVL04-16-M02 (AY072303) as a template"

1..1614

/gene="At3g14610"

/note="synonym: MIE1.11"

1..1539

/gene="At3g14610"

/product="cytochrome P450"

/protein\_id="AA03051.1"

/db\_xref="GI:28059362"

/translation="MSFSVVAALPLVAVVLTWRIYKWKIKEMLESSLKQGLT

GPYTPVLDIKRVNMMKARSKPINTVDITPLPLPALMLSHKGTFFIIGPL

PTIYTPNPDQIKRVNMMKARSKPINTVDITPLPLPALMLSHKGTFFIIGPL

EKIKRMIPAPYHCSRYVCOMKELPTDKSPLEUVPMVUNMTADYVSHRBSYK

EGQRIPOGLSLALIAQAKSYIPGSPFYPTKSRKMAIDREVDVILGIVSKK

KAREGAPANDDLGLILSNSESQGNMSVEDWKRCKLYPEAGQSTTVLLVMT

VLSHODQAFARAEVYQVLGNKNDPESLNNKVMIMFNEVRLYPRVAOLKEV

VNKMKLGELTLPAGIOIYLPILVORDELWDGDAAPKPRFPDGLSKATKQVF

FPFMGPRICIGQFALFKXAMALILQKSFELSPSYHAPQVTMTTQFGAHLI

LHKL"

## ORIGIN

Query Match 20.4%; Score 318.8; DB 8; Length 1614;

Best Local Similarity 52.8%; Pred. No. 3.7e-80;

Matches 761; Conservative 0; Mismatches 672; Indels 9; Gaps 3;

QY 90 GTTATGTGAGACCAAGAAAGATTGAAGAACTTCTTAACAAGAAATCGAGTCC 149  
 DB 78 GGTCTGATTAACCAAGATGCTTGAAGTCTTTGAAGAGAGGGCTTACAGGAAC 137  
 QY 150 TCTTATCATTTTTCATCGGAATGTTAAGAACTTTTGAATGATGCTTAAGCTTC 209  
 DB 138 TCTTCACTCTCTCTCGAGATTAAGAAAGAAATGTTGATGATGATGAAAGCAG 197  
 QY 210 TTTTCATCCATGCTTCTCTCAATATTTCTTCAAGTTCCTCTTTTACATCA 269  
 DB 198 ATCTAAACCATCAATGATGAAGATGATACACCCACGCTCTCTCTGCTTAA 257  
 QY 270 CTGAGAAAGAAATCTACGCTGATCATTTCTGTTGGTTCGCTCACTTCGGATTAAC 329  
 DB 258 GATGCTCATTTCTACGAGAAAGACTTTTCAATGATGATGACCACTTCCACGATTT 317  
 QY 330 GGTAGCCATCTCATATTGATCAAGAGATCTTCTTAAGTCTGATTTCTACAGAGAA 389  
 DB 318 GATTAAGATCTTGAAGATCAAGAGAGCTTTTAATTAAGTCAACGATTTGAAGAAC 377  
 QY 390 TGAAGTCACTCTTGGTTAAACAATTGAAGGCGATGACTACTTAAGTCAAGAGTGA 449

DB 378 TTTTACATTCCTTTGATCAGATTGTTAGCAG--TGAGCTTGCAAGTTTCAAGGAGA 434  
 QY 450 AAAATGGGTCATCATCAAAAATCATATAGCCCTACTTTTCATATGAGAGATTAAGTT 509  
 DB 435 TAAATGGGAGATCAGAGAGATCATCAACCGGCTTTTCACTCGAAAAATCAAGAA 494  
 QY 510 GCTTGATCAAGTGTGTTGAAGAGTGAATGATGATGATGATGATGATGATGATGAT 569  
 DB 495 CATGATCCCTGGCTTCAACATTTGAGAGAGATGTTGTCTCAATGGAGAGATTAAT 554  
 QY 570 ATCAGAAAACG--GTGAAGTTGAGTAGATGCTATGATGATGATGATGATGATGATGAT 626  
 DB 555 TACAGATTAAGATGCTCTTGAAGTCAATGTTGGCTTGGCTTGGCTTGGCTTGGCTTGG 614  
 QY 627 AATGTTATTAAGTAACAAGCTTTTGAAGTATGATGATGATGATGATGATGATGATGAT 686  
 DB 615 GATGTCATCTCACTACTGCTTTTGAAGTATTAAGAGAGAGAGAGAGAGATTAATTA 674  
 QY 687 ACTTCAAGCTCAACAAATGCTTCTTGTGCTGAAGCTTTCAAAAAGCTTTCATTCCTGG 746  
 DB 675 ACTTCAAGAGAGATTTGCTGAGCTTATGAGCAAGCTTTTAAAGAAATCTTACATCCTGG 734  
 QY 747 CTATGATTTTTTCCGATAGAGAGAAATTTGAAGTCTCGAAGTTAGCAAGAGATTAAG 806  
 DB 735 ATCGAGGTTTATCCCAAG 794  
 QY 807 GAAGTCGTTGTTGAAGCTGATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 866  
 DB 795 CGTAAATTTTGAAGAGATTTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 854  
 QY 867 AATATGTAAG 923  
 DB 855 AATGATATTTGTTGGAGATATTTGCTTAATCAATTCAGAGAAATCTTCAAGAGATAG 914  
 QY 924 TGTGACGTTGAGAGCAATTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 983  
 DB 915 AATGAGCGTAGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 974  
 QY 984 AACTTCTAATGCTGAG 1043  
 DB 975 AACTTCAATGCTTTGTTGAG 1034  
 QY 1044 CAAGACGATGATGAG 1103  
 DB 1035 TCGGCAAG 1094  
 QY 1104 TGTGTTAGCTTAAAG 1163  
 DB 1095 CTTTAAACAGCTTAAAG 1154  
 QY 1164 AATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1223  
 DB 1155 GGTGCTGAGCTTAAAG 1214  
 QY 1224 AATGAG 1283  
 DB 1215 AACTGAAATTTCAAAATTTACTTACCACTATTTCTTTCAGAGAGAGAGAGAGAGAGAG 1274  
 QY 1284 TAATGACGTGAAGAAATTCATTCAGCTGAGTTTGGAGTGGAGAGAGAGAGAGAGAGAG 1343  
 DB 1275 CGATGATGACAGAGAGATTTAAACCGAGCGGTCAAGAGAGAGAGAGAGAGAGAGAGAG 1334  
 QY 1344 AAGCCGAGTGGCTTCAATCGGTTGGCTCGAGAGTTGCTGATGATGATGATGATGATGATG 1403  
 DB 1335 GAACAGAGCTCTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1394  
 QY 1404 TGTATTAAGTCAAGCAAAATTTGACCTGCTGATGATGATGATGATGATGATGATGATGAT 1463  
 DB 1395 TGTATGTTGAG 1454  
 QY 1464 GGTGCTTCTTATGAG 1523  
 DB 1455 CTCTCTCTTCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1514

QY 1524 AC 1525  
 Db 1515 TC 1516  
 RESULT 6  
 AY072303  
 LOCUS  
 DEFINITION Arabidopsis thaliana cytochrome P450 (At3g14610; MIE1.11) mRNA,  
 complete cds.  
 ACCESSION AY072303.1 GI:18252154  
 VERSION AY072303  
 KEYWORDS FL1 cDNA.  
 SOURCE Arabidopsis thaliana (thale cress)  
 ORGANISM Arabidopsis thaliana  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
 1 (bases 1 to 1786)  
 Nguyen,M., Karlin-Neumann,G., Southwick,A., Lam,B., Miranda,M.,  
 Palm,C.J., Bowser,L., Jones,T., Bam,J., Carninci,P., Chen,H.,  
 Cheuk,R., Chung,M.K., Hayashizaki,Y., Ishida,J., Kawai,J.,  
 Kawano,H., Kim,C., Lin,U., Liu,S.X., Natsumaka,M., Pham,P.K.,  
 Sakano,H., Sakurai,T., Saitou,M., Seki,M., Shim,P., Yamada,K.,  
 Shinozaki,K., Ecker,J., Theologis,A. and Davis,R.W.  
 Direct Submission  
 Submitted (02-JAN-2002) DNA Sequencing and Technology Center,  
 Stanford University, 855 California Avenue, Palo Alto, CA 94304,  
 USA  
 e-mail for correspondence: arab@sequence.stanford.edu  
 COMMENT  
 RIKEN Genomic Sciences Center (GSC) members carried out the  
 collection and clustering of the RAFL cDNAs (RAFL cDNA: 'RIKEN  
 Arabidopsis Full-length cDNA'): Seki,M., Natsumaka,M., Ishida,J.,  
 Saitou,M., Kawai,J., Sakurai,T., Carninci,P., Kawai,J.,  
 Hayashizaki,Y. and Shinozaki,K.  
 The Saik, Stanford, PGSC (SSP) Consortium members carried out the  
 sequencing and annotation of the RAFL cDNAs: Nguyen,M.,  
 Southwick,A., Karlin-Neumann,G., Lam,B., Miranda,M., Palm,C.J.,  
 Bowser,L., Jones,T., Bam,J., Chen,H., Cheuk,R., Chung,M.K.,  
 Kim,C., Lin,U., Liu,S.X., Pham,P.K., Sakano,H., Shim,P.,  
 Yamada,K., Ecker,J., Theologis,A. and Davis,R.W.  
 Nguyen,M. (SSP/Stanford) and Seki,M. (RIKEN GSC) contributed  
 equally to this work. Shinozaki,K. (RIKEN GSC) and Davis,R.W.  
 (SSP/Stanford) contributed equally to this work as PIs.  
 FEATURES  
 source  
 1..1786  
 /organism="Arabidopsis thaliana"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:3702"  
 /chromosome="3"  
 /clone="RAF104-16-M02"  
 /ecotype="Columbia"  
 /note="This clone is in pBluescript"  
 1..1786  
 /gene="AC3g14610; MIE1.11"  
 100..1638  
 /gene="AC3g14610; MIE1.11"  
 /product="cytochrome P450"  
 /protein\_id="AA161910.1"  
 /db\_xref="GI:18252154"  
 /translation="MSFSVVALPVLVAVVLTWIRIVKWKIRKLSLKEQGIT  
 GPTPLVDIKAVDMMEARKSPINVDITPRLIPALRKMSGKFFIIGL  
 PTVITNPOIKVENKRVNDFKASTPLIRLLAGGLASVGDKMSRRIRINAFHL  
 EIKRMIPAPFYHCSEVYVCOMKELFTDESPLVAVMLVMTADVISHAFSSYK  
 EGORITQLOGLAEHLIAQAFKSYIPGSRFPTKSNRMKALIDEVVLIRGISVKE  
 KAREGEPAINDLLGILSSESSESGMSVDEWMECKFYRAGQVTLVILWMT  
 VLSHQDQARAREVWQVLENNKPMDESLNKLKVTIMLFENEVLRYPVVALQKV  
 VKEMELGELTIPAGIQIYLPILVIRDTLMGDDAADFKEFRFDGSKATKQVSP

ORIGIN  
 LHLK"  
 FPFMGPRICIGGNFAMLEAKMALLIQKSFELSPSYVHAPQVWTRPQFALHI  
 Query Match 20.4%; Score 318.8; DB 8; Length 1786;  
 Best Local Similarity 52.8%; Pred. No. 3.8e-80;  
 Matches 761; Conservative 0; Mismatches 672; Indels 9; Gaps 3;  
 QY 90 GTTATGGGAGACCAAGAAAGATTAAGAACTTCTTAACAAGAAATTCAGAGTCC 149  
 Db 177 GGTCTGATTAACCAAGAAAGATTCGTTGAGATCTTTGAAAAGACAGAGCTTACCGGAAC 236  
 QY 150 TCCTTATCATTTCTTCATCGGAAATGTTAAGAACTTTGGATATGATGCTTAACTTC 209  
 Db 237 TCCTTACACTCCTTCGTCGAGATTAAGAAAGAAATGTTGATATGATGAGAACCGAG 296  
 QY 210 TTCTCATCTATGCTCTTCTCTCAATATCTTCCAGAGTCTCTCTTTTACATCA 269  
 Db 297 ATCTAAACCCATCATGATTAACGATGATATCAACCCACGCTCTCTCTCTTCCCTTAA 356  
 QY 270 CTGGAGAAATCTACGAGTCTACATTTCTGTTGGTGGTCCAACTTCGCTTAC 329  
 Db 357 GATGCTCAATTCACCGAAAGAACTTCTCATATGATGAGCACTTCCAAAGATGT 416  
 QY 330 GGTACCGATCTCTGATTTGATCAGAGAGATCTTCTAAGTCTGAGTTCTACGAGAA 389  
 Db 417 GATTAACGATCTTACGAGATCAAGAGAGCTTTTAATTAAGTCAAGACATTTAGAAAGC 476  
 QY 390 TGAAGTCAACCTTGGTGAACAACCTGAAGGAGATGAGATCTTATGTCGAAAGTGA 449  
 Db 477 TTCTACATCTCTTGTGATGATGATGATGAGAG---TGGCTTGCAAGTTACAGGAGA 533  
 QY 450 AAAATGGCTCATCATGAAATATCATTAACCTTCTCATATGAGAACTTAACT 509  
 Db 534 TAAATGGCGAGTCAAGAGAGATCAACCCGCTTTCACCTGCAAAAATCAAGAA 553  
 QY 510 GCTTGTACCACTTGGTGAAGAGTGAAGTGAATGATGATGATGATGATGATGATGAT 559  
 Db 554 CATGATCCCTGCTTCTCACTTTTGGAGAGAGTGTCTGTCATAGGAGAAAGCTATT 653  
 QY 570 ATCAGAAACG---GTGAAGTGAAGTGAATGATGATGATGATGATGATGATGATGAT 626  
 Db 654 TACAGATTAAGATCGCTCTTGAAGTGAATGATGATGATGATGATGATGATGATGAT 713  
 QY 627 AGATGTTTATGATGAAGAGCTTTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 686  
 Db 714 GATGTCATCTCACTACTGCTTTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 773  
 QY 687 ACTTCAGCTCAACAAATGCTTCTTGTGCTGAAGCTTTTCAAAAAGTCTTCACTCTCG 746  
 Db 774 ACTTACAGGGAATGGCTGAGCTTATGAGCAAGCTTTTAAAGAACTTACATCTCTCG 833  
 QY 747 CTATGATTTTTTCCGACAAAGGGAATTTGAATCTTCGGAAGTTTACAAAGGATTAAG 806  
 Db 834 ATCGAGGTTTACCAAGAAAGAGCAATGAAGATGAAGCAATGAAGATGAAGATGAAG 893  
 QY 807 GAAGTCGTTGTAAGTGAAGTGAAGCGCGGAGCAAAACGCTTATGAATGAAGAGGGA 866  
 Db 894 CGTATATTTGAAGATATTTGAGGAAGAAAGGGAAGAAAGCGAGAGAGTGAAGACGAG 953  
 QY 867 AAGATGTAAGAGCGCGGCGGAGAGATTTGTTGGATTAATGA---TTACGCAAGAA 923  
 Db 954 AAACGATGATTTGTTGGGGAATATTCCTTGAATTCATTCAGAGAAATTCAGAGAAATGG 1013  
 QY 924 TGTGACGTTTCAGACATTTGTGAGAGAGTGAAGATTTTCTTCTGCGCGGAAACAGAC 983  
 Db 1014 AATGACCGTGAAGATGATGAAGAAAGATGAAGATGAAGATGATGAAGATGAAGATGAAG 1073  
 QY 984 AACTTCAATCTGCTGAAGTGAAGCAATCTTGTATATCATGACACCGGAGTGGCGAGGC 1043  
 Db 1074 AACTTCAATCTTGTGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1133  
 QY 1044 CAAAGACGATGAGAGTCTCAGAGGTCTGCGGCTCAGATGATGCTTCAACAAAGAGCA 1103

[illegible]

COMMENT  
RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RafL cDNAs (RafL cDNA : 'RIKEN Arabidopsis Full-length cDNA') : Seki, M., Narusaka, M., Ishida, J., Satou, M., Kamiya, A., Sakurai, T., Carninci, P., Kawai, J., Hayashizaki, Y. and Shinozaki, K.

The Salk, Stanford, PGSC (SSP) Consortium members carried out the sequencing and annotation of the RafL cDNAs: Kim, C.J., Chen, H., Cheuk, R., Meyers, M.C., Shinn, P., Banb, J., Bowser, L., Chang, E., Dale, J.M., Goldsmith, A.D., Jones, T., Karlin-Neumann, G., Lam, B., Lee, J.M., Lin, J., Miranda, M., Nguyen, M., Onodera, C.S., Palm, C.J., Quach, H.L., Southwick, A., Tang, C.C., Toriumi, M., Wu, H.C., Yamada, K., Yamamura, Y., Yu, G., Yu, S., Davis, R.W., Theologis, A., and Ecker, J.R.

Kim, C.J. (SSP/Salk) and Seki, M. (RIKEN GSC) contributed equally to this work. Shinozaki, K. (RIKEN GSC) and Ecker, J.R. (SSP/Salk) contributed equally to this work as pIs.

FEATURES  
source  
1..1548  
/organism="Arabidopsis thaliana"  
/mol\_type="mRNA"  
/db\_xref="taxon:3702"  
/chromosome="3"  
/clone="U13574"  
/ecotype="Columbia"  
/note="this clone is in pUNI 51"  
1..1548  
/note="putative cytochrome P450"  
/codon\_start=1  
/product="AT3g14620/ME1.12"  
/protein\_id="P138603.1"  
/db\_xref="gi:17386114"  
/translation="MSDPTKSAVAVAIAVVVVTVTVIMWIKGLNAVMLPKNKNAVLYLK  
ROGSGVFTFLVGDIDIKRQASMTQEKSRPINLDYTPRARNPLDQYVXKHGKTSYV  
WSPFLKSTVTKPPEHIDKQVLRVYDFKPRVPIYELPATGVALTEGKMKRHTIA  
BSFLEKTKIMIPAFVSCSEMTSKWEKTSQSSNEIDWPVYIGDLSVYSTA  
SGYSGGRIFELQEOGRVILALSLAPFQGRFLPTSNLMQINQVAKSVLRRE  
IKRQYRGDTEAGPKNDLIGLILLESNGSHGMSIEDVVEBCILFFHQAEATVILVAV  
TMIMSHQKQODQARBEILKVIQGNKKKNFDSLKTMMSIMLNEVLYLPQGLILDL  
RVWETKTCGMDMLPGAGAVTVPLWHRDPELMEDVHEENPFRADGSKTKNK  
VSFLPFGKPRFCQGNFALMEKMLVILIQFSELSPSVTHA.PHIVLILHPQGR  
PLIFHML"

ORIGIN  
Query Match: 18.7%; Score 291.6; DB 8; Length 1548;  
Best Local Similarity 52.2%; Pred. No. 2.5e-72;  
Matches 770; Conservative 0; Mismatches 669; Indels 15; Gaps 5.

70 ATAGTAAAGGTATGCTCTCTGTTATGTTGAGAGACCAAGAAAGTTGAAGAATCTTCTCT 129  
73 ATATGAAAGGCTGAATAGTAGCTGGTGTAGACCAAGAAAGAAATAGGCTTTCCTAAG 132  
130 AAACAAGGAATCGAGGCTCCCTCTTATCATTTCTTCACGGAATGTTAAAGAACTGTT 189  
133 AGCAAGAGTCTCTCGAAGCTCTTTCACCTTTTGTGGAGACATTAAGAGGAGGCC 192  
190 GAATGATGCTTAAAGCTTCTTCTCATCTTATGCTTCTCTCACAAATATCTTCTCTAA 249  
193 AGTATGATGAGCAAGAAAGTCCAGACCCCATCATATTAATGATGATTAACGCAATGCT 252  
250 GTTCTCTCTTTTACCATCATCGAGAAATAATACAGTGTCTACATTTCTGGTTTGGTTC 309  
253 GTCATGCTCTTATTAACACAGACCGTTAAGATGATGGAAGAACATTTACATGCTGGAG 312  
310 GATCCAACTTCCGTTAACGTTACGTTACCGATCTCTGATTTGATCAGAGAGATCTTCTTA 369  
313 GAGACCTTACGAGATGATGATGTTGAGCAAGAACCAAGAACATCAAGAAAGTCTTAAACA 372  
370 TCGAGTGTTCAGAGAAAGATGAAGTCAACCTTGTGGTTAAACAATCTTGAAGCGATGA 429  
373 GTTACGATTTCCCAAGCCGCGACATGACCAATCGTT---GAGCTCTTCGGAATCGGA 429

QY 430 CTACTAGTCTCAAAAGTGAATAATGGGCTCATCATCGAAATCATTTAGCCCTACTTTT 489  
DB 430 GTCCGCGTGTATGAGAGAGAAATGGTCTTAAGCAGAGAAATTTAAACCTTTCTTTT 489  
QY 490 CATATGAGAAATCTTAAGTTGCTGTATACCAATGTGTGTTGAAGAGTGAAGTGAATGATG 549  
DB 490 CATCTTGAAGAAAGCTCAAGATTAATGATACCTGCATTTACAGAGCTGACGCGAGATGATA 549  
QY 550 GATAAATGCTCCGATTAAGTTATCGAAAGCGTGA---GTGAAGTGAATGCTTATGAG 606  
DB 550 AGTAATGGGAGAAATGTGTGACCGACAGAGATCATGAGATGAGATGATGTTGGCCA 609  
QY 607 TGGTTTCAGATTTTGAAGTGAAGATGTTATTAAGAACAGCTTTTGAAGTGAAGTGAAG 666  
DB 610 TATCTTGAAGATTTAATCATGAGATGATTTTCAAGTACTGCAATTTGGTATGATCTGAG 669  
QY 667 GATGTCGAGCAGTTTTCGACTTCAAGCTTCAACAAATGCTTTCTTGTCTGAGACTTTT 726  
DB 670 GAAGTGAAGAAATCTTGAAGCTTCAAGAAAGACAGCGCGAGAGTGTAAAGCTTTT 729  
QY 727 CAAAAGTCTTCACTTCCGCTATGATGATTTTTCGACAAAGAGGAAATTTGAAGTCTCG 786  
DB 730 GAGTGGCTTTTATCCCTGAGATGAGGTTTCTACCAAGACAAATTTGAGAGATGAGA 789  
QY 787 AAGTTAGACAAAGAGATTAAGAGATGCTGTTGTAAGCTGATAGACGCGCGAGACAAAC 846  
DB 790 CAAATTAACAAAGA---AGTGAATCTAGACTGAGAGATCATGAAAGACAGAGG 846  
QY 847 GCTATAGATGAGAAAGGAGAAATGTAAGAGCCGCGCGAGAGATTTGTTGGATTA 906  
DB 847 GGTATGATATACAGGAGAGCTCCAAAGATGATCTGTTAGGA---TACTTTGGAATCG 903  
QY 907 AAGATTCAGGCAAAAGATGATGACGTTGAGAGATTTGAGAGATTAAGCTTTTTC 966  
DB 904 AATTTGAGATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 963  
QY 967 TTGCGCGGAGAAACAGACAACTTTTATCTGCTGACGTGAGACCAATCTTGTATCAGT 1026  
DB 964 TTCCGCGGCGCAAGAAACACCGCGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1023  
QY 1027 CACCGCGAGTGGCAGGCGCAAGACATGATGAGGCTCAGGCTCGCGCTCAGCTGAT 1086  
DB 1024 CACCAAAATGCGAGAGATCACTGATGAGAAAGATTTAAAGTCAATTTGGGAGAAAC 1083  
QY 1087 GTCCCTACCAAGACATGCTGTTAAGCTTAAGCTTAAGCTTAAGCTTAAGCTTAAGCT 1146  
DB 1084 AATCCATTTTGAAGCACTCTCTGCACTCAAAACATGACATGATCTTGAAGAAAGTT 1143  
QY 1147 TTAAGTGTATTCACCAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1206  
DB 1144 TTGAAGTGTATTCACCAAGATGATGATGATGATGATGATGATGATGATGATGATGAT 1203  
QY 1207 GAGAGGTAC---AAATCCCATGATGATGATGATGATGATGATGATGATGATGATGAT 1263  
DB 1204 GGTGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1263  
QY 1264 CATGACCAAGCAATTTGGGATGATGATGATGATGATGATGATGATGATGATGATGAT 1323  
DB 1264 CGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1323  
QY 1324 GAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1383  
DB 1324 GGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1383  
QY 1384 ACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1443  
DB 1384 TTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1443  
QY 1444 CAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1503  
DB 1444 CAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1503  
QY 1504 CTTATCTCTCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1537

DB 1504 CTTATCTCTCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1537  
RESULT 8  
LOCUS AY052208  
DEFINITION Arabidopsis thaliana AT3g14620/M1E1\_12 mRNA, complete cds.  
ACCESSION AY052208.1  
VERSION GI:15529168  
KEYWORDS FTI CDNA.  
ORGANISM Arabidopsis thaliana (thale cress)  
REFERENCE  
AUTHORS Cheuk, R., Chen, H., Kim, C. J., Koesema, E., Meyers, M. C., Ban, J., Bower, L., Carninci, P., Dale, J. M., Goldsmith, A. D., Hayashizaki, Y., Ishida, J., Jiang, P. X., Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J., Lam, B., Lee, J. M., Lin, J., Liu, S. X., Miranda, M., Narusaka, M., Nguyen, M., Onodera, C. S., Palm, C. J., Pham, P. K., Quach, H. L., Sakurai, T., Satou, M., Seki, M., Southwick, A., Tang, C. C., Toriumi, M., Yamada, K., Yamamura, Y., Yu, G., Yu, S., Shinzaki, K., Davis, R. W., Theologis, A. and Ecker, J. R.  
TITLE Arabidopsis CDNA clones  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 1784)  
AUTHORS Cheuk, R., Chen, H., Kim, C. J., Koesema, E., Meyers, M. C., Ban, J., Bower, L., Carninci, P., Dale, J. M., Goldsmith, A. D., Hayashizaki, Y., Ishida, J., Jiang, P. X., Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J., Lam, B., Lee, J. M., Lin, J., Liu, S. X., Miranda, M., Narusaka, M., Nguyen, M., Onodera, C. S., Palm, C. J., Pham, P. K., Quach, H. L., Sakurai, T., Satou, M., Seki, M., Southwick, A., Tang, C. C., Toriumi, M., Yamada, K., Yamamura, Y., Yu, G., Yu, S., Shinzaki, K., Davis, R. W., Theologis, A. and Ecker, J. R.  
TITLE Direct Submission  
JOURNAL Submitted (16-AUG-2001) Salk Institute Genomic Analysis Laboratory (SIGAL), Plant Biology Laboratory, The Salk Institute for Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037, USA  
COMMENT  
The Salk, Stanford, PGEC (SSP) Consortium members carried out the sequencing and annotation of the RAPL cDNAs: Cheuk, R., Chen, H., Kim, C. J., Koesema, E., Meyers, M. C., Shin, P., Ban, J., Bower, L., Dale, J. M., Goldsmith, A. D., Jiang, P. X., Jones, T., Karlin-Neumann, G., Lam, B., Lee, J. M., Lin, J., Liu, S. X., Miranda, M., Nguyen, M., Onodera, C. S., Palm, C. J., Pham, P. K., Quach, H. L., Southwick, A., Tang, C. C., Toriumi, M., Yamada, K., Yamamura, Y., Yu, G., Yu, S., Davis, R. W., Theologis, A., and Ecker, J. R.  
FEATURES  
source  
1..1784  
/organism="Arabidopsis thaliana"  
/mol\_type="mRNA"  
/db\_xref="taxon:3702"  
/chromosome="3"  
/clone="RAF06-77-B22(R13574)"  
/ecotype="Columbia"  
1..53  
54..1601  
/note="putative cytochrome P450"  
/codon\_start=1  
/product="AT3g14620/IE1\_12"  
/protein\_id="AAK97679.1"



/db\_xref="GI:15529169"  
 /translation="MSDTKISAVAAVAVVTVTVTWKGLNAMLPRKNEAYLK  
 RQSGTPEFVLDIKREASNVEOKSRINLTDYTRVAPLLQOVKHGKSTVM  
 WMGPJASVITKPEHIDVLRNRYDFKPPHIVLEFVTVGLVPEGKMSKRIIN  
 GSFLKELKIMIPAFVSESEMSKWEKVTGEGSENEIDVPYGLDLSVI SRTAF  
 GSYEGRKRIPELOEGRVLEKALFLAPGRLPTKMLRMOIKVEYSLREI  
 IMKROMDGCEAKNDLIGILLESNGHMSIEDVEECLEFPAECFTATLIVL  
 TMILSHQKQODARBEILKVI GKNNKRNPDLSLTKMSITLNEYRLIPGILLG  
 RIVSEKTRLEGDMLTLPGRQVVI PVMWEDBELWDEHNEPERADGSKATKQ  
 VSLPFGMPKPCPGQNFALMEKMLVLLQPSFELSPTTHAPVTLTLHPQFA  
 PLIFHML"  
 1602..1784

Query Match 18.6% Score 290.6; DB 8; Length 1784;  
 Best Local Similarity 52.0%; Pred. No. 5e-72;  
 Matches 775; Conservative 0; Mismatches 699; Indels 15; Gaps 5;

70 ATAGTGAAGGATGCTCTGTTATGTTGAGACCAAGAAAGTTGAAGACATTTCT 129  
 126 ATATGGAAGGCTGAGATGATGCTTGAACCAAGAAAGTTGAAGCTTATTAAG 185  
 130 AAACAGAAATTCAGAGCTCTCTTATCATTTCTTATGGAAGTTGAAGACTTGT 189  
 186 AGACAAAGGCTCTCTGGAACCTCTTACCTTTCTTGTGGAACATTTAAAGGAGACC 245  
 190 GGAATGATGCTTAAAGCTCTTCTCATCTTATGCTCTCTCAAAATATCTTCTAGA 249  
 246 AGTATGTTGAGCAAGAAAGTCCAGACCATCAATCTAATGATGATTAACAGCATGT 305  
 250 GTTCTCTTTTACCATCTGAGCAAAATCTAAGTGTCTATCTTGTGTTGTTTC 309  
 306 GTCATGCTCTGATPACACACACCGTTAAGATCAAGAGACATTTACATGAGATG 365  
 310 GGTCCAACTTCCGGTTAACGGTACCGATCTGATTTGATGAGAGATCTTCTTAAG 369  
 366 GGAACCTATAGGAGGTGATGTTGACAAACCAAGAACATCAAGATTTCTTAAGACA 425  
 370 TCTGATTTCTAGCAAGAAATGAAAGTCAACCTTTGGTTAAACAATTGAAGCGATGA 429  
 426 GTTTACGATTTCCGAGCCGCGCATGACCCCAATGTT--GAGCTCTTCCGACATGA 482  
 430 CTACTTACTTCAAAAGTGAAGAAATGGGCTCATCATCAAGAAATCTTAGCCCTACTT 489  
 483 GTCCGGCTGATGAGAGGAGAAATGCTTAAACACAGGAAGTTTAAACCTCTTTT 542  
 490 CATATGAGGATCTTAAGTCTTGTTCAGATGTGTGTAAGAGTGTGATGATGTTG 549  
 543 CATCTTGAAGGCTCAAGATTTATGATCTGCACTTCAAGAGGCTGACGCAATGATA 602  
 550 GATTAATGCTCCGATTAAGTTATCAAGAAACGATGA--GTTAGGATAGATCTTATGAG 606  
 603 AGTAAATGAGGAAGTGTGATGACCAAGATCAATCAAGATGATGATGATTTGGCCA 652  
 607 TGTGTTCAAGATTTGACGAAAGATTTATGAGAAACGCTTTGGAAGTGTATGA 666  
 663 TATCTTGAAGATTTAACTGATGATGATTTATTCAGTCTGATGATGATGATGATG 722  
 667 GATGTCGAGCAGTTTTCAGCTTCAAGCTCAACAAATGCTTTTGTGCTGAAGCTTT 726  
 723 GAAAGTAAAGAAATTTTGAAGCTTCAAGAAACAAAGCGGCGAGTGTATTAAGCTCTT 782  
 727 CAAGAAATCTCATTCCTGCTATAGATTTTTCACCAAGAGGAAATTTGAAGTCTGG 786  
 783 GATGTCCTTTTATCCCTGAGATGATGATTTCTTACCAAGAAACAAATTTGAGATGAGA 842  
 787 AAGTTGAACAAGAGATTAAGAGTCTGTTGAAGTGTATGAGCGCGGAGACAAAC 846  
 843 CAATTAATCAAGAA--AGTAAATCTAGACTGAGAGATCTCATGAAAGAGACAGAG 899  
 847 GGTATGATGAGAGAGGAGAAATGTAAGAGCGCGCGGCAAGATTTGTTGGATTA 906  
 900 GGTATGATTAAGGAGGAGCTTCCAAAGATGATCTGTTAGAA--TACTTTTGAATG 956

907 ATGATTCAGGCAAGAAATGTGACGGTTCAAGACATTTGTGAGAGATGTAAAGCTTTTC 966  
 957 AATTTCTGAGATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1016  
 967 TTGCGCGGGAACAGACAACTTTAATCTGCTGACGAGACGACCATCTTGTATTCATG 1026  
 1017 TTGCGCGGCAAGAAACCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1076  
 1027 CACCGGAGTGGCAGGCGCAAGACGATGAGGCTCTGAGGCTTGGCGCTCACTGAT 1086  
 1077 CACCAAAATGCGAGATCAAGCTAGAGAAAGATATTAAGCTCATTTGGGAAAGAACAC 1136  
 1087 GTCCTACCAAGACCAATGCTGATTAAGCTTAAACGTTAGATGATCTTGAACGAGCT 1146  
 1137 AAACCTAATTTTGAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1196  
 1147 TTAAGTGTATTCACCAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1206  
 1197 TTGAGGTTATCCACCAAGGATCTACTTGTGCTGAAACCTGAGAAAGAAACAAAGCTT 1256  
 1207 GGAGGATCA--AAATCCATGAGCAGGAGCTTAAATCCATCATGAGGCTCAT 1263  
 1257 GGTGAAGCATGACATCAAGGATGATGATGATGATGATGATGATGATGATGATGATGAT 1316  
 1264 CATGACCAAGCCTATTTGGGATGATGATGATGATGATGATGATGATGATGATGATGAT 1323  
 1317 CGTGAATCCGAGCTCTGGGATGATGATGATGATGATGATGATGATGATGATGATGAT 1376  
 1324 GGAGTCCGCGGCTGCAAAACACCGCTTGGCTTATACCTTTGGCTCGGAGTTCTG 1383  
 1377 GGTATCTAAAGCAGCAAGAAACCAAGTTCTGTTCTTCCGTTGGATGAGGAGCAAGAA 1436  
 1384 ACATCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1443  
 1437 TTCTGCTGCGCAGAAATTTGCTTATGATGATGATGATGATGATGATGATGATGATGAT 1496  
 1444 CAAGCTTCACTTCACTTGGCTCTCACTTATCAAGATGATGATGATGATGATGATGATG 1503  
 1497 CAGATATCTCTTGTGAGTGTCTCTTCTGATCACTGACCATCATGAGGTTGTACA 1556  
 1504 CTTTATCTCAACATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1552  
 1557 CTTCACTCTCACTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1605

RESULT 9  
 AK120018  
 LOCUS  
 DEFINITION  
 Oryza sativa (japonica cultivar-group) cDNA clone:002-186-B11, full insert sequence.  
 AK120018  
 ACCESSION  
 AK120018.1 GI:37989641  
 VERSION  
 FLI CDNA; oligo capping.  
 KEYWORDS  
 Oryza sativa (japonica cultivar-group)  
 Oryza sativa (japonica cultivar-group)  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzaceae; Oryza.  
 1  
 REFERENCE  
 AUTHORS  
 The Rice Full-Length cDNA Consortium, National Institute of Agricultural Sciences Rice Full-Length cDNA Project Team; Kikuchi, S., Satoh, K., Nagata, T., Kawagashira, N., Doi, K., Kishimoto, N., Yazaki, J., Ishikawa, M., Yamada, H., Ooka, K., Hotta, I., Kojima, K., Nemiki, T., Ohneda, E., Yahagi, W., Suzuki, K., Li, C., Ohtsuki, K., Shibasaki, T., Foundation of Advancement of International Science Genome Sequencing & Analysis Group; Otsu, Y., Murakami, K., Iida, Y., Sugano, S., Fujimura, T., Suzuki, Y., Tsunoda, Y., Kurokawa, T., Kodama, T., Masuda, H., Kobayashi, M., Xie, Q., Lu, M., Narikawa, R., Sugiyama, A., Mizuno, K., Yokomizo, S., Nikiura, J., Ikeda, R., Ishibiki, J., Kawamata, M., Yoshimura, A., Miura, J., Kusumegi, T., Oka, M., Ryu, R., Ueda, M., Matsubara, K., RIKEN, Kawai, J., Carninci, P., Adachi, J., Aizawa, K., Arakawa, T., Fukuda, S.,



Harada, A., Hasinidume, W., Hayatsu, N., Imocami, K., Ishii, Y., Itoh, M., Kawaga, I., Komodo, S., Komoh, H., Miyazaki, A., Otsuki, Y., Ota, Y., Saito, R., Saeki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T., Yoshino, M., and Hayashizaki, Y. Collection, mapping, and annotation of over 28,000 cDNA clones from

Journal of Science 301 (5631), 376-379 (2003)

PUBMED 12869764

## AUTHORS

Aodachi, J., Ueda, K., S., Akimura, I., Atakawa, I., Cahnline, P., Doi, K.,  
 Fujimura, T., Fukuda, S., Hanagata, T., Hara, A., Hashitume, W.,  
 Hayashida, K., Hayashizaki, Y., Hayatsu, N., Himemoto, K., Hiraoka, T.,  
 Hori, F., Hotta, I., Iida, J., Iida, Y., Ikeda, R., Imamura, K.,  
 Imecanci, K., Ishihiki, J., Ishii, Y., Ishikawa, M., Itoh, M., Kagawa, I.,  
 Kangawa, S., Katoh, H., Kawagashira, N., Kawai, J., Kawamata, M.,  
 Kikuchi, S., Kishikawa-Hirozane, T., Kishimoto, N., Kobayashi, N.,  
 Kodama, T., Kojima, K., Kojima, T., Kusumegi, T., Kuwano, H., Kouda, H.,  
 Koya, S., Kurihara, C., Kurotaki, T., Kusumegi, T., Li, C., Lu, M.,  
 Masuda, H., Matsubara, K., Matsuyama, T., Mitsu, J., Miyazaki, S.,  
 Mizuno, K., Murakami, K., Murata, M., Nagata, T., Nakahama, Y.,  
 Nakamura, M., Namiki, T., Narioka, R., Nikiura, J., Nishi, K.,  
 Nomura, K., Nunasaki, R., Ohmeda, E., Ohno, M., Ohtsuki, K., Oka, M.,  
 Ooka, H., Ootao, N., Ota, Y., Otomo, Y., Ryu, R., Satoh, H., Sakai, C.,  
 Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Sato, K., Satoh, K.,  
 Shibata, K., Shingawa, A., Shiraki, T., Shinbiki, T., Sogabe, Y.,  
 Sugano, S., Sugiyama, A., Suzuki, K., Suzuki, Y., Tagami, M.,  
 Tagami, Takada, Y., Tagawa, A., Takahashi, F., Takaki-Akahira, S.,  
 Tanaka, T., Tomaru, A., Toyota, T., Tsunoda, Y., Ueda, M., Waki, K.,  
 Xie, D., Yahagi, M., Yamada, H., Yamamoto, M., Yasunishi, A., Yazaki, J.,  
 Yokomizo, S., and Yoshimura, A.,  
 collection, mapping, and annotation of 28K full-length cDNA clones  
 from japonica rice

## REFERENCE

TITL E

1

## FEATURES

— 100 —



Db		777	GAGCGCATGAGAGATGTCTACATTCCAGGATGGAGCTACTTGCCAAACAAGAGAACAAG	836
Oy		778	AAGTCTCGGAAGTTTAACAAGAGATTAAGAAAGTCGTGTTGAAGCTGATGAGCGCGG	837
Dp		837	AGGATGAAAAGAAATTAACAAAGAGATCAAGATATGCTAAGGTTCAATCACAAGAGA	896
Oy		838	AGA CAAAA CGCTTAGATGAGAAAGGGAAGAATGTAAGAGCGCGCGGGAAGATTGG	897
Dp		897	ATGAGAGGCTTTGAAGGCTGAGAGGCAGGTAGAGATGACTTGGTGGGAATGTTGTTGAA	956
Oy		898	T-----TGGAGTTAATGATTCAAGCCAAGATGTGACGGTTGAG	936
Dp		957	TCAAACTTTCANAGAAATTCAAAAGAACAGAAAACAAGAGATGGTGAATGTCATCAAT	1016
Oy		937	GACATTGTGAGAGAGTGTAAAAGCTTTTTCTTCGCCGGGAAAACAGCAACTTCTATCTG	996
Dp		1017	GANGTATATGAGAGAGTGAATAATGTTCTACTTTGCTGCTGTCAGAAACTACTGGAAGTTTA	1076
Oy		997	CTGACGTGGAACCAACCATTTGCTATCCATGACACCAGAGTGGCAGGCGCAAAGACGTAT	1056
Dp		1077	CTGACATGAGCACCACTTTATTGACAAGACCCCTTAGATGGCAAAAGCAGACAGAGA	1136
Oy		1057	GAGGTCCTCAGGGTCTCGCGGCTCAAGTATGTCCCTCAAGAGACCATGTGCTTAAGCTT	1116
Dp		1137	GAA GTTCTCCAAAGCTTTGGC--AAGAAATAACCTGATGTTGAAGGCTTAATCACTCG	1192
Oy		1117	AAAACGTGAGATATATCTTGAACGAGCTTTAAGCTGTATATCCACATTAAGACTAG	1176
Dp		1194	AAATATGTGTCTATGATCTTGTACGAGGCTCTTAAGTTGTACCAACAGTATGATCTA	1255
Oy		1177	ATTGACGCGCTMAATCGAGTGTGAAGCTAGAGAGGTACAAAATCCCATGTGGCAAGG	1238
Dp		1254	ACAAAGATTGTCCACAAGAGACAAAGATTGAGGTGTACAAATCCCTGACAGGAACA	1313
Oy		1237	CTTCTATGCCCAATCA TACGGGTGCATATACCAACCCATTTGGGGTATAGACGTGAAC	1296
Dp		1314	GTATGTGTGCCAGAGTATATGCTTCAAGAGAAGAGACATTTGGGAGAGATGTCATG	1373
Oy		1297	GAA TTCATCA TCACGCTCGGTTTGCGAGTGAAGTGC GCGGTGCTGCAACACCCCGTTGGC	1356
Dp		1374	GAA TTC AACCCATAGATGATTTGTGATGAGATTTGCCAATCAACAAAGAACATGTA CA	1433
Oy		1357	TTCA TCCGTTGGCGCTCGGAGTGTGTATCAATGATGATGATGAGATCTGTGTAATCTGAG	1416
Dp		1434	TATTTGCCATTACGCTGCGGACCTTAGGGTTGTCTTGGCCAAAATTGCACTTCTGCA	1493
Oy		1417	GCCAAATTAACA CTGCTGTGTAATGATCCACGCTTCACTTTCACTGGCTCCTACTAT	1476
Dp		1494	GCAAAATTAAGATTGGCAATGATTTTACAAAGCTTCAAGTTATGATGCTCCATCTAT	1553
Oy		1477	CACATGACACATACGCTCCTTATGTTGCTTATATCTCAAGATGATGACCAATCACTCT	1536
Dp		1554	GTTCAATGCTTCCTTTACTACTTCTCAAGATTCAACCCAGTTGGTCTCATGTCACTCAC	1613
Oy		1537	CGGA 1540	
Dp		1614	AMGA 1617	
<hr/>				
RESULT 11				
CTRP450B				
LOCUS	CTRP450B	1633 bp	mRNA	linear
DEFINITION	Catharanthus roseus Cp3 cytochrome 450 (CYP72C) mRNA, 3' end.			
VERSION	L19075			
KEYWORDS	L19075.1 GI:404689			
SOURCE	cytochrome 450.			
ORGANISM	Catharanthus roseus (Madagascar periwinkle)			
	Catharanthus roseus			
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			
	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;			
	asterids; lamiales; Gentianales; Apocynaceae; Rauvolfiaceae;			
	Vincaceae; Catharanthus			
REFERENCE				
	1 (bases 1 to 1633)			

**AUTHORS** Mangold, U., Eichel, J., Bartschauer, A., Lanz, T., Kaiser, T., Spangenberg, G., Weick-Reichert, D. and Schroeder, J.P. (1997) Gene and cDNA for plant cytochrome P450 proteins (CYP72 family) from *Catharanthus roseus*, and transgenic expression in tobacco and *Arabidopsis thaliana*  
**JOURNAL** Plant Sci. 96, 129-136 (1994)  
**REFERENCE** 2 (bases 1 to 1633)  
**AUTHORS** Joachim Schroeder.  
**TITLE** Direct Submission.  
**JOURNAL** Submitted (30-SEP-1993) by Joachim Schroeder, Biologie II, Univ. of Freiburg, Schanzlestr. 1, Freiburg, Germany D-79104  
**COMMENT** Original source text: *Catharanthus roseus* (strain Cp3) cDNA to mRNA.

**FEATURES**  
 source  
 Location/Qualifiers  
 1..1633  
 /organism="Catharanthus roseus"  
 /mol\_type="mRNA"  
 /strain="Cp3"  
 /db\_xref="taxon:4058"  
 /cisseq\_id="NM1149"  
 1..1633  
 /gene="CYP72C"  
 <1..1553  
 /gene="CYP72C"  
 /note="putative"  
 /codon\_start=3  
 /product="cytochrome P450"  
 /protein\_id="AA017746.1"  
 /db\_xref="GI:404650"  
 /translation="1AATIFALVAMAMRVLDAMFTPKRIEKRLIQGFRGNPYRFIVDVSFGSKHQEALSNPEKENDNDIVRLMPHINHIKYGNUSFTPMKRIPIRHVMHPELIEVESGKMSKYQKNDVNPVYKLLGVSGFSGAKSKHRIISPAFLLEKIKMLPAPIICHDMLTWELKAEKSGSEVDIFPPFDVSPVISKVAFGTYDEGKIKRILKEIMLTDICMDKDYVILPMSYLVPRKMKKEINKETIDMLRFIINRKMALKAABEGEDILGVILESESIQLOKQGRKPGMTINDVIECKLPFAQDETGVLLTWTIILSKHFERQERARREVLQATGRKKEPERFENMLKYVSMILYVRLVLPVLDLKIHHDTKLGEPITPAQVWMLPTVMLHREKSIWGDADAEFENPRADSVNATKANTYIYPSWBRVFLQNFALLOAKGLAMILLQRFEPVADSVYAAPEPTILTVOPQGSVITIKLERONF"

**ORIGIN**  
 Query Match 18.0%; Score 281.2; DB 8; Length 1633;  
 Best Local Similarity 51.1%; Pred. No. 2,56-69;  
 Matches 757; Conservative 0; Mismatches 686; Indels 27; Gaps 3

QY	83	TGTCCTGTTATGSGTGAAGACCAAGAAAGATTGAAGACATTTCTTAAACAAGCAATTC	142
DB	52	TGATTTGGGCAATGGTTTACTCTTAAAGAGATCGAAGAACGCTTAAGGACGCAAGGTTT	111
QY	143	GAGTCCCTCATATTTCTTCAACGGAATGTTAAGAATTGTTGATGATGCTTA	202
DB	112	GAGGAATCTTATGATTTCTTGTTGAGATGTTAAGAGAGTGGAAAAATGCATCAG	171
QY	203	AAGCTTCTTCATCTCATGCTTTCTCTCACAATATTTCTCTAAGATTTCTCTTTT	262
DB	172	AAGCTTGTCTAACCCCAAGAGATTCACCAACGATATGTTCTCGCTCAATGCCAATA	231
QY	263	ACCATCACTGAGAAAAATCTACGGTGTACATTTCTGTTTGGTTCCGTCCAACTTCC	322
DB	232	TTAACACACCACTAAAACTTACGGTAGAATTCATTTATCATGAGTGGAGAGATCCAA	291
QY	323	GGTTAAGGTAGCCGATCTGATTTGATCAAGAGATCTCTTAAGTGGAGTCTTAG	382
DB	292	GGATTCATGTTATGAACTGAATCTTAAAGAAAGATTAAGCACTCAAGAAATACC	351
QY	383	AGAGAA---TGAAGTCAACCTTTGTTAAACAATTGAAGGAGATGACTTATGTC	439
DB	352	AAAGAACTTATGTTTACAATCCCTTTGTTAAGTCTCTCACTGSGAGTTGAAGCT	411
QY	440	TCAAAGTAAATAAGGGGTCTATCGAAAAATCATTAAGCCCTACTTTTATATGAGA	499
DB	412	TTGAAGGGCAAAATGTCCTCAACACGAAATAATCATTTCCCTGCTTCACTCTTGAGA	471

QY 500 ATCTAGATTGCTTGTACAGATTGTGTGTAAGATGTGATATGTTGATTAATGCT 559  
 DB 472 AACTTAAGTCAATGTTGCCGGCTTTGGCATTATGCTATCATGACATGTTGACGAAATGGG 531  
 QY 560 CCGATTAAGTTATCAGAAAACGGTAAAGTTAGGTATGCTATGATGGTTTCAATTT 619  
 DB 532 AGAAATTAGCAGAAAAGAAAGATGCCATGAATGATATCTTCCCAATTTGATGTTT 591  
 QY 620 TGAATGAAGTATTTATTAAGTAAACAGCTTTGGAAGTATGATGAAGATGATGATGAG 679  
 DB 592 TGACAAAGTATGTAATTTCAAGGTTGCAATTTGGTGCATATGATTAAGAGGCAAAA 651  
 QY 680 TTTTGCATCTCAAGCTCAACAATGCTTTTGTGCTGAAGCTTTTCAAAAAGTCTTCA 739  
 DB 652 TCTTCAGACTATTTGAAAAGAACTCATGATCTCATATGATTCATGAGATGCTTACA 711  
 QY 740 TTCTGGCTATATGATTTTTCGACAAAGAGGAATTTGAAGTCTCGAAGTTAGCAAG 799  
 DB 712 TCCAGGGGTGAGCTACTTGCACAAAGAGAAAGAGATGAAGAAATTAACAAAG 771  
 QY 800 AGATPAGAAAGTGGTTGTGAAGCTATAGAGCGCGAGACAAAACGCTATATGATGAG 859  
 DB 772 AATTCACAGATATGCTAAGTTATCATCATCAAGAAATGAAGGCTTTGAAGGCTGAG 831  
 QY 860 AAGGGAAGAAATGTAAGAGCCGCGCG-----CGAAGATTGT 898  
 DB 832 AACCAAGTGAAGATGATTTGTTGGAGTCTGTTGGAATCAAAATTCAAGAAATTCAGA 891  
 QY 899 TGGGATTAATGATTCAGGCAAAAGATGACGTTGAGACATATGAGAGAGTATAAA 958  
 DB 892 AACAGGAACAGAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 951  
 QY 959 GCTTTTCTTCCGCGGAAAACAACAATTATCTGCTGACGTGACGACATCTTGC 1018  
 DB 952 TGTCTACTTGTGCTGCTCAAGAAACCCGAGTTTGTCTCATGACGACACCATTTAT 1011  
 QY 1019 TATCATGACACCCGAGTGGCAGGCGCAACAGCTATGAGGCTCAGGATCTGCGGCT 1078  
 DB 1012 TGACCAAGCATCCGAGTGGCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1070  
 QY 1079 CACGTGATGTCCTTACCAAGAGCCATGCTTAACTTAACGTTAAGTATGATCTTGA 1138  
 DB 1071 --AAGAACAGCCTGAATTTGAAGCTTGAATCATCTTAATATGTTTCTATGATCTTAT 1128  
 QY 1139 ACGAGCTTTAAGTTGATTCACCAATAGTACGTGATGATGACCGCGCTAATGGGATG 1198  
 DB 1129 ACGAGGTTCTAAGTTGATTCACCGCAGTTATGATGATGATGATGATGATGATGAT 1188  
 QY 1199 TGAAGCTAGAGGCTACAAATCCATGCTGAGCGAGGCTTCTAATCCATCATATGAG 1258  
 DB 1189 CAAAGTTAGGCGCTTACCAATCCCTGACAGAAACAAATATGTTGCCGACGATATGC 1248  
 QY 1259 TCCATCATGACCAAGCCATTTGGGCTAATGACGTGAACGAATTCATCCAGCTCGGTTG 1318  
 DB 1249 TTCAATAGAGAGAGGAGCTTTGGGAGAGAGATGCAATGAAATTCACCCCAATGATTTG 1308  
 QY 1319 CGGATGAGAGCGCGGCTGCTGCCAAACACCCGTTGCTCATACGTTTGGCTCGAG 1378  
 DB 1309 CTGATGAGTTGCCAATGCCACCAAGAAACATATGTCATATTTGCCATTCAGTTGGGAG 1368  
 QY 1379 TTCTGATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1438  
 DB 1369 CTAGAGTTTGGCTCGGCAAAACCTTTCATTCAGCAAGAAATTAAGGTTGGCAATGA 1428  
 QY 1439 TGAATCAAGCTTCAACCTTCACTGATGATGATGATGATGATGATGATGATGATGAT 1498  
 DB 1429 TTTTCAAGGCTTCAAAATTTGATGATGATGATGATGATGATGATGATGATGATGAT 1488  
 QY 1499 TGTGCTTATCTTCAACATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1540  
 DB 1489 TTAAGTTCAACCTCAGTTGTTCTCATGATGATGATGATGATGATGATGATGATGAT 1530

## RESULT 12

AY096748 1570 bp mRNA linear PLN 18-SEP-2002  
 LOCUS Arabidopsis thaliana putative cytochrome P450 protein (At3g14690)  
 DEFINITION mRNA, complete cds.  
 ACCESSION AY096748.1 GI:20465786  
 VERSION  
 KEYWORDS  
 SOURCE  
 ORGANISM

## REFERENCE

1 (bases 1 to 1570)  
 Yamada, K., Ban, H., Chan, M.M., Chang, C.H., Chang, E., Dale, J.M., Deng, J.M., Goldsmith, A.D., Lee, J.M., Onodera, C.S., Quach, H.L., Tang, C.C., Toriumi, M., Wu, H.C., Yamamura, Y., Yu, G., Bowser, L., Carninci, P., Chen, H., Cheuk, R., Hayashizaki, Y., Ishida, J., Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J., Kim, C., Lam, B., Lin, J., Meyers, M.C., Miranda, M., Narusaka, M., Nguyen, M., Palm, C.J., Sakurai, T., Satou, M., Seki, M., Shim, P., Southwick, A., Shinozaki, K., Davis, R.W., Ecker, J.R. and Theologis, A.  
 Arabidopsis Open Reading Frame (ORF) Clones  
 Unpublished  
 2 (bases 1 to 1570)  
 Yamada, K., Ban, H., Chan, M.M., Chang, C.H., Chang, E., Dale, J.M., Deng, J.M., Goldsmith, A.D., Lee, J.M., Onodera, C.S., Quach, H.L., Tang, C.C., Toriumi, M., Wu, H.C., Yamamura, Y., Yu, G., Bowser, L., Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J., Kim, C., Lam, B., Lin, J., Meyers, M.C., Miranda, M., Narusaka, M., Nguyen, M., Palm, C.J., Sakurai, T., Satou, M., Seki, M., Shim, P., Southwick, A., Shinozaki, K., Davis, R.W., Ecker, J.R. and Theologis, A.  
 Arabidopsis Open Reading Frame (ORF) Clones  
 Unpublished

## TITLE

## JOURNAL

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## REFERENCE

## AUTHORS

## COMMENT

The Salk, Stanford, PGC (SSP) Consortium members constructed and sequenced the PUN1 (ORF) clones using the RFL CDNA. Yamada, K., Ban, H., Chan, M.M., Chang, C.H., Chang, E., Dale, J.M., Deng, J.M., Goldsmith, A.D., Lee, J.M., Onodera, C.S., Quach, H.L., Tang, C.C., Toriumi, M., Wu, H.C., Yamamura, Y., Yu, G., Bowser, L., Chen, H., Cheuk, R., Kamiya, A., Karlin-Neumann, G., Kawai, J., Kim, C., Lam, B., Lin, J., Meyers, M.C., Miranda, M., Narusaka, M., Nguyen, M., Palm, C.J., Sakurai, T., Satou, M., Seki, M., Shim, P., Southwick, A., Shinozaki, K., Davis, R.W., Ecker, J.R. and Theologis, A.  
 Direct Submission  
 Submitted (16-APR-2002) Plant Gene Expression Center, 800 Buchanan Street, Albany, CA 94710, USA  
 The RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RFL CDNA. RFL CDNA: 'RIKEN Arabidopsis Full-length CDNA': Seki, M., Narusaka, M., Ishida, J., Satou, M., Kamiya, A., Sakurai, T., Carninci, P., Kawai, J., Hayashizaki, Y. and Shinozaki, K.

## FEATURES

source  
 1. 1570  
 /organism="Arabidopsis thaliana"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:3702"  
 /chromosome="3"  
 /clone="F11126"  
 /ecotype="Columbia"  
 /note="This clone is in pUN1 51."  
 1. 1570  
 /gene="At3g14690"  
 1. 1539  
 /gene="At3g14690"  
 /codon\_start=1  
 /evidence=experimental  
 /product="putative cytochrome P450 protein"

Annotation is based on the January 2002 version of the Arabidopsis genome submitted to GenBank.  
 Location/Qualifiers  
 1. 1570  
 /organism="Arabidopsis thaliana"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:3702"  
 /chromosome="3"  
 /clone="F11126"  
 /ecotype="Columbia"  
 /note="This clone is in pUN1 51."  
 1. 1570  
 /gene="At3g14690"  
 1. 1539  
 /gene="At3g14690"  
 /codon\_start=1  
 /evidence=experimental  
 /product="putative cytochrome P450 protein"

/protein\_id="AA020382.1"  
 /db\_xref="gi:20465787"  
 /translation="MEISVASVTISVYLAVSWIWRLLQWVFKRMLKHLRQGL  
 AGTYPLVDLKNFTLSPASRSLKLTDDISPRVYPLQMTYGRITFTWGP  
 IPTITMDPEQIKENFKYDFOKEHPLATIIAGLAVDGDMAKRLINPWF  
 IEXIKMVPAPHOSCRVYGEWDQVSDKSCSEVWPELVMTADVSRFAFSY  
 KEGORIPELOAELOLIIOAFKAFIPGYSLPTKSRKMAAREIOVILGIVKR  
 LBARAGBAPSDDLIGLIESNLOTGEMSTEDLMESCKLFYPAGORTSVLWV  
 MYLSQHDMABARAREVKOYRGDKEDPAQINOLKVMILYEVRLVPTQLTRA  
 IHKELKSDLTLPQGVQSLPIILVCHDIELMKNDAEFNPDPKGLSKATKSOVSF  
 FPFAPGPRICIGONFALLEAKMMALILRFSEISIVHAYITITHPQAGLI  
 MHKL"  
 1540..1570  
 /gene="At3g14690"

## ORIGIN

Query Match 17.7%; Score 276.4; DB 8; Length 1570;  
 Best Local Similarity 51.3%; Pred. No. 6,1e-68;  
 Matches 746; Conservative 0; Mismatches 696; Indels 12; Gaps 4;

QY 90 GTTATGTGGAGACCAAGAAATTGAAGACATTCTCTPAACAAGAAATTCGAGTCC 149  
 DB 81 GGTGGTTTAAAGCAAGATGCTTGAACATTAAGTGAAGACAAAGTCTTCGCGAAC 140  
 QY 150 TCCCTATCATTTCTTCATCGAATGTAAAGAACTTGTGAATGATGCTTAAAGCTTC 209  
 DB 141 ACCTTAACAGCCTTCAAGTCGGGATTTGAAGAAATTTCCACATGTTGTGCGAGCGAAC 200  
 QY 210 TTCTATCTTATGCTTCTCTCAAAATTTCTCCAGAGTTCTCTTTTACATCAATCA 269  
 DB 201 GTCCAAACCCCTCAACTAAGCATGATCTCACTCTGTCGTGCTTATCCCTTGA 260  
 QY 270 CTGAGAAAATCTACGGGTGATCAATTTCTGTTGTTTCCGTCAACTTCCGGTTAC 329  
 DB 261 AATGTTCAAGACTTACGGAAGACCTTACTTAACATGTTTGAAGCTTACCAACATCAC 320  
 QY 330 GGTAGCCGATCTGATTTGATGAGAAATCTTCTTAAGTCTGAAGTTCAAGGAAGA 389  
 DB 321 AATTAATGATCTTGCAATCAAGAAATGTTCAAAAGTTTATGATTTCCAAAGCC 380  
 QY 390 TGAAGTCACTCTTGGTTTAAACAATGAAAGCGATGACTTACTAGTCTCAAGGTGA 449  
 DB 381 GCATACATTTCTTTGGCCA---CAATTATAGCTTAAGAGACTCCGATTAATGATGTA 437  
 QY 450 AAAATGGGCTCATCTGAAATGAAATCATTAAGCCCTTACTTATATGAGAAATCTTAAAGTT 509  
 DB 438 TAAATGGGCAAAACCGAAGATCATCAACCGGCTTTTCAATATAGAGAAATCAAGAA 497  
 QY 510 GCTTGACCAAGTTGTTGAAGATGATCTGATATGATGATTAATGATGTCGATTAAGTT 569  
 DB 498 TATGATCTGCTTCCACCAAGCTGTAAGAAATTTTGGCCAAATGGAATCTAGT 557  
 QY 558 ATCGAATAAAGGTGCTCATGAGGTAGACGTTTGGCGCTGTGATGATAGACTGC 617  
 DB 627 AGATGTTTATTAAGAAAGCTTTTGAAGTGAATGAAGTGTGAGAGAGTTTTCG 686  
 QY 618 AGATGATCTCTCGTACGCAATTTGGCGAGCTTCAAAAGAGGCGAAGAGATTTTGA 677  
 DB 687 ACTCAAGCTCAACAATGCTTCTTGTGCTGAGCTTTCAAAAGTCTTCAATCTCG 746  
 QY 678 GCTCCAGAGAACTGACAGAGTATCATTTCAAGTTTATAGAAAGCTTTCATCCCTCG 737  
 DB 747 CTATAGATTTTTCGACAAAGAGAAATTTGAAGTCTGGAAGTTTGAACAAGAGATAG 806  
 QY 738 ATATAGTATCTCCCAACAAAAGTACAGAAATGAAGACGACGACAGAAATCTCA 797  
 DB 807 GAAGTGGTTTGAAGCTGATAGAGCGCGAGACAAAACGCTATATGATGAGAAAGGGA 866  
 QY 798 AGTTATCTAAGAGGATCTTAACAAAAGTTTAAAGGCGACAGAAAGCTGTGAAGCACCC 857  
 DB 867 AGATGTAAAGAGCCGCGGCAAGATTTGTTGGATTAATGATTTACGCAAAAGATGT 926

DB 858 AAGCGAGATTTGTGGGAATTAATCTTGTGATGAATTTGAGGCAACGAAAGAAACGG 917  
 QY 927 GAGCGT---CAGACATTTGAGAGAGTAAAGACTTTTCTTCCCGGGAACAGAC 983  
 DB 918 AATGATACCGAGATCTGATGAGAGAGTGCAGTGGTTCTATTTTCCCGGCGAAGAC 977  
 QY 984 AACTTAAATCTGTCAGCGTGAACGACCATTTGCTATCCATCCACCCGAGTGGACAGC 1043  
 DB 978 AACATCAGTACTTCTGTTTGAACATGATGCTGTTAAGCAACACCAAGATTGGACAGC 1037  
 QY 1044 CAAAGCCGATGATAGATCTCTCAAGGCTGCGGCTGACGATGTCCTTCCCAAGACCA 1103  
 DB 1038 TGTGACAGAGAAAGTGAAGCAAGTGTGGCGA---TAAAGAACTGATGCAGAAAG 1094  
 QY 1104 TGTGTTAAGCTTAAACGTTGATGATGATCTTGAACGAGCTTTAAGTTGATCCACC 1163  
 DB 1095 ACTCAACAGCTCAAGATTAATGATGATGATATTAATGAGTCTTAAAGCTTATCTTCC 1154  
 QY 1164 AATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1223  
 DB 1155 AGTAAACCAAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1214  
 QY 1224 ATGAGGACAGGAGCTTCAATCCCAATCATAGCGGTCCATCATGACCAAGCAATTTGGGG 1283  
 DB 1215 AGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1274  
 QY 1284 TATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1343  
 DB 1275 AAACGACGACAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1334  
 QY 1344 AACCCCGTGGCTTCAATACCGTTGGCTGCGAGTTCGATGATGATGATGATGATGATGAT 1403  
 DB 1335 GAGCAAGTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1394  
 QY 1404 TGTATATCTTCAAGCCAAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1463  
 DB 1395 TGCCCTGTTAAGGCAAAATGCAATGATGATGATGATGATGATGATGATGATGATGAT 1454  
 QY 1464 GCTCTCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1523  
 DB 1455 TTCTCTCTCTATGTTTCAAGCAATGATGATGATGATGATGATGATGATGATGATGAT 1514  
 QY 1524 ACCAATCACTTCC 1537  
 DB 1515 TCAAGTTATCATGC 1528

RESULT 13  
 AY050827  
 LOCUS  
 DEFINITION  
 mRNA, complete cds.  
 ACCESSION  
 VERSION  
 KEYWORDS  
 SOURCE  
 ORGANISM

REFERENCE  
 AUTHORS  
 TITLE  
 JOURNAL  
 REFERENCE  
 AUTHORS

1692 bp mRNA linear PLN 18-SEP-2002  
 Arabidopsis thaliana putative cytochrome P450 protein (At3g14690)  
 Yamaoka, K., Liu, S. X., Sakano, H., Pham, P. K., Ban, J., Chung, M. K.,  
 Goldsmith, A. D., Lee, J. M., Quach, H. L., Toriumi, M., Yu, G., Bowser, L.,  
 Carninci, P., Chen, H., Cheuk, R., Hayashizaki, Y., Ishida, J.,  
 Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J., Kim, C., Lam, B.,  
 Lin, J., Miranda, M., Narusaka, M., Nguyen, M., Palm, C. J., Sakurai, T.,  
 Satou, M., Seki, M., Shim, P., Southwick, A., Shinzaki, K.,  
 Davis, R. W., Ecker, J. R. and Theologis, A.  
 Arabidopsis Full Length cDNA Clones  
 Yamada, K., Liu, S. X., Pham, P. K., Ban, J., Banjo, F., Dale, J. M.,

TITLE  
JOURNAL  
COMMENT

Goldsmith, A.D., Jiang, P.X., Lee, J.M., Onodera, C.S., Quach, H.L., Tang, C.C., Toriumi, M., Yamamura, Y., Yu, G., Yu, S., Bowser, L., Carrinci, P., Chen, H., Cheuk, R., Hayashizaki, Y., Ishida, J., Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J., Kim, C., Koeseema, E., Lam, B., Lin, J., Meyers, M.C., Miranda, M., Narusaka, M., Nguyen, M., Palm, C.J., Sakurai, T., Satou, M., Seki, M., Shim, P., Southwick, A., Tracy, S.E., Shinohara, K., Davis, R.M., Ecker, J.R. and Theologis, A.

Direct Submission  
Submitted (08-AUG-2001) Plant Gene Expression Center, 800 Buchanan Street, Albany, CA 94710, USA  
RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RAPL cDNAs (RAPL cDNA: 'RIKEN Arabidopsis Full-length cDNA'): Seki, M., Narusaka, M., Ishida, J., Satou, M., Kamiya, A., Sakurai, T., Carrinci, P., Kawai, J., Hayashizaki, Y. and Shinohara, K.

The Salk, Stanford, PGRS (SSP) Consortium members carried out the sequencing and annotation of the RAPL cDNAs: Yamada, K., Liu, S.X., Pham, P.K., Ban, J., Banno, F., Dale, J.M., Goldsmith, A.D., Jiang, P.X., Lee, J.M., Onodera, C.S., Quach, H.L., Tang, C.C., Toriumi, M., Yamamura, Y., Yu, G., Yu, S., Bowser, L., Chen, H., Cheuk, R., Jones, T., Karlin-Neumann, G., Kim, C., Koeseema, E., Lam, B., Lin, J., Meyers, M.C., Miranda, M., Nguyen, M., Palm, C.J., Shim, P., Southwick, A., Tracy, S.E., Davis, R.M., Ecker, J.R. and Theologis, A.

Yamada, K. (SSP/PGRS) and Seki, M. (RIKEN GSC) contributed equally to this work. Shinohara, K. (RIKEN GSC) and Theologis, A. (SSP/PGRS) contributed equally to this work as PIs.

Annotation is based on the January 2002 version of the Arabidopsis genome submitted to GenBank.

FEATURES  
source

Location/Qualifiers

1..1692  
/organism="Arabidopsis thaliana"

/db\_xref="taxon:3702"

/chromosome="3"

/clone="RAF108-11-K16 (R11126)"

/ecotype="Columbia"

/note="This clone is in a modified pBluescript vector (pLC-1) as a BamHI/XhoI insert."

1..1692  
/gene="At3g14690"

1..31  
/gene="At3g14690"

32..1570  
/gene="At3g14690"

/codon\_start=1

/evidence=experimental

/product="putative cytochrome P450 protein"

/protein\_id="AAK2762.1"

/db\_xref="gi:15292787"

/translation="MEISVASVTISVLA VSWIWRITQWVFKRNL EHYLRQGL AGTYPLVGLKNTMLSEARSKPLKLDISPRVYVPLQMTKRYGTYFTMFG IPTTIMDPOIKREKRVYDPOKHPPLATIIAGLNVGDKAKGRILINPAFI IEXIKMVPAPHOSCEVEMDQVDSKSSCEVAVGVLSTADVISIATAGSSV KESORIELOEALOLIOAFRAPIRGSYLPTKSNRMAAARLIVIIAGVNR IARABGABSDDLGITLESNI RPTGNSMSTEDMECKI FYRAGCTTSVLTWT VVLSQRDWRARAREKOVFGKSDAGLNQLVNTMIIIVELRIIPVTLQTRA IHKRLKGLDITLPGVQISLPIILVQHDIELWGNDAENPDRFGDSKATKQVSP FPFAMGPRICTGQNFALLAKAMALILRFSPFSPSYVAHYVITIHPOFAGOLI MHKL"

1571..1692  
/gene="At3g14690"

1676  
/gene="At3g14690"

/note="not present in genomic sequence"

## ORIGIN

Query Match 17.7%; Score 276.4; DB 8; Length 1692;  
Best Local Similarity 51.3%; Pred. No. 6.2e-68;  
Matches 746; Conservative 0; Mismatches 696; Indels 12; Gaps 4;

Qy 90 GTTATGTCGAGACCAAGAAAGATTGAAGAACTTTCTTAACAGAAATTGAGATCC 149  
Db 112 GATTGGTTTAAAGCAAGATGCTGAGCATTAAGTAAAGAGAGCTTCCGGAAC 171  
Qy 150 TCTTATCATCTTCTTCATCGAAATGTTAAAGAACTTTGGAATGATGCTTAAAGCTTC 209  
Db 172 ACCTTACAGCCCTTAGTCGAGATTGAAGAAATTTCAACATGTTGCGAGGCAAG 231  
Qy 210 TTTCTATCTATAGCTTCTCTCAATATTTCTTCAAGTCTCTCTTTTACATCA 269  
Db 232 GTCCAAACCCCTCAAACTTAAGAGATGATCACTCTGATGCTGATCTTATCCCTTGA 291  
Qy 270 CTGGAAGAAATCTACGGTGTGATCTTCTGTTTGTGCTCAACTTCCGTTAAC 329  
Db 292 AATGTTCAAGACTTACGAGAGACTTACTTACATGTTTGAACCTATCCACATCAG 351  
Qy 330 GGTAGCCATCTCTATTTGATCAAGAGATCTTCTAATCTGAGTTCTAGAGAGAA 389  
Db 352 AATATGATCTCTGCAAAATCAAGAGATGTTACAAAGTTATGATTTCAAAAGCC 411  
Qy 390 TGAAGCTACCTTGTGTTAAACAATTGAAGGCGATGACTATTAGTCTAAAGTGA 449  
Db 412 GCATCAATTTCTTTGGCCA---CAATTAATGTTAAAGACTCGTATATTATGATGTA 468  
Qy 450 AAAATGGCTATCATCAAAATATTAAGCCCTTCAATATGAGAAATCTTAAGT 509  
Db 469 TAAATGGGAAACCGAGAAATCATACCCGGCTTTCATATAGAGAGATCAAGAA 528  
Qy 510 GCTTGTACAGTTGTGTTAAAGATGATGATGATGATGATGATGATGATGATGAT 569  
Db 529 TATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 588  
Qy 570 ATCAAGAAACGGTGA---AGTTAGGTTAGATGATGATGATGATGATGATGATGAT 626  
Db 589 ATCGATTAAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 648  
Qy 627 AGATGTTATTAAGTAAACAGCTTTTGAAGTATGATGATGATGATGATGATGATGAT 686  
Db 649 AATATGATCTCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 708  
Qy 687 ACTTGAAGTCAAGAAATGCTTTTGTGCTGAAAGCTTTCAAAAGTCTTCTCTG 746  
Db 709 GCTCAAGAGATCAAGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 768  
Qy 747 CTATGATTTTTCGCAAGAGAGAAATTTGAAGTCTGGAAGTGAACAAGAGATTAAG 806  
Db 769 ATATGATTAATCTCCCAAGAAAGTAAAGAGATGATGATGATGATGATGATGATGAT 828  
Qy 807 GAAAGCTGTTTGAAGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 866  
Db 829 AGTTATTAAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 888  
Qy 867 AGAATGTAAGAGCCGCGGCGGAAGATTTGTTGGATTAATGATTAAGGCAAGATGT 926  
Db 889 AAGCAGCATTTGCTGGAAATTAATCTTGAATCAATTTAGGCAAGAGAGAAACG 948  
Qy 927 GACGGT---CAGCAATTTGAGAGAGTGAAGTAAAGTCTTTCTGCGCGGAAACAG 983  
Db 949 AATGATACCGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1008  
Qy 984 AACTTAAATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1043  
Db 1009 AACATCAAGTACTTCTGTTTGAAGAGAGTGTGATGATGATGATGATGATGATGATGAT 1068  
Qy 1044 CAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1103  
Db 1069 TGTGATCAAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1125  
Qy 1104 TGTGATCAAGTAAAGCTTGAAGTATGATGATGATGATGATGATGATGATGATGATGAT 1163  
Db 1126 ACTCAAGAGCTCAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1185





Db	Accession	Version	Keywords	Source	Organism	Reference	Authors
OY	AK101545	LOCUS DEFINITION	AK101545	AK101545	AK101545	RESULT 15	
OY	1039	CAGGCCAAAGCAGCGTATGAGGTCTCTGAGGGTCTGCGGCTGACGATGATCCCTTCAAG	1098				
Db	1066	CAGAGCGAGCAGAGAGAGAGATGCTTACACACTTTGG---AAGAACCAACACAGACAT	1122				
OY	1099	GACCATGTCCTTAAGCTTAAACGTTAGATGATCTTGAACAGAGCTTTAAGTTGAT	1158				
Db	1123	GATAGCTTAAGTCGTGTGAAGTTGTATCAATGATTTGTACACAGGTTCTTGAAGTTGAT	1182				
OY	1159	CCACCAATAGTAGCTACGATTCGATTCGACGCGCTAAATCGATCTGAGGTAGAGGATACAA	1218				
Db	1183	CCGCGCTGTGTCTTCTTGACACGAGAACATACAGAGAAATGAGTCTGGCGGATCAAA	1242				
OY	1219	ATCCCATGTGACACGAGAGCTTTATTCCTCAATCATACGCGTCCATCTATACCAACCAT	1278				
Db	1243	TATCCGCTGAAAGTACCCCTTATGTGTGCCATTTTATTTTACATCATATCCCATATT	1302				
OY	1279	TGGGATATAGCTGACGATTCGAATTCATATCAGCTCGGTTTCCGATGAGATGCGCGTCT	1338				
Db	1303	TGGGAGAAAATATGACAGATGAATTCAATCCAGAGAGTTTCTGATGGATCTCCACAGCA	1362				
OY	1339	GCCAAACACCCCGTTGCTTCAATACCGTTTGGCTCGAGTTGTATCATGCAATTTGTGAG	1398				
Db	1363	ACGAAATATGACAGCTCTTTCTTTCATTTGGATGGGTTCCCGAATTCGATCGCGCTAG	1422				
OY	1399	AATCTGCTATCTACTCAGGCGCAATTTGACATCGCTGTATGATTCACAGCTTCACTTT	1458				
Db	1423	AATCTGCACTATTTGAAACCAATATGCTATCTGTACATCTTCAAGGGTTCTCTTT	1482				
OY	1459	CACTTGCTCTTACTTATATGAGATGACGACTTACGCTTATGTGCTTTATCTTCACAT	1518				
Db	1483	GAGCTTTCACCATGATCGATCCACGACACATTTCACTGTGATTAATCTTCCACCCACAGAT	1542				
OY	1519	GGTGACCAATCACTCTTCCGAGATT	1544				
Db	1543	GGTGACCAATTAAGTTGAAGAAAT	1568				
RESULT 15	AK101545	1903 bp	MRNA	linear	PLN 24-JUL-2003		
LOCUS	Oryza sativa (japonica cultivar-group)	cdna clone:U03048K24	full				
DEFINITION	insert sequence.						
ACCESSION	AK101545	1	GI:32986754				
VERSION	AK101545.1						
KEYWORDS	Full cDNA; CAP trapper.						
SOURCE	Oryza sativa (japonica cultivar-group)						
ORGANISM	Oryza sativa (japonica cultivar-group)						
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzaceae; Oryza.						
AUTHORS	1. The Rice Full-length cDNA Consortium, National Institute of Agricultural Sciences Rice Full-length cDNA Project Team; Kikuchi, S., Satoh, K., Nagata, T., Kawagashira, N., Doi, K., Kishimoto, N., Yazaki, J., Ishikawa, M., Yamada, H., Ooka, H., Hotta, I., Kojima, K., Namiki, T., Ohneda, E., Yanagi, M., Suzuki, K., Li, C., Ohtsuki, K., Shishiki, T., Foundation of Advancement of International Science Genome Sequencing & Analysis Group; Otsuo, Y., Murakami, K., Iida, Y., Sugano, S., Fujimura, T., Suzuki, Y., Tsunoda, Y., Kurokawa, T., Kodama, T., Masuda, H., Kobayashi, M., Xie, Q., Lu, M., Narikawa, R., Sugiyama, A., Mizuno, K., Yokomizo, S., Nishizawa, J., Ikeda, R., Ishibiki, J., Kawamura, M., Yoshimura, A., Miura, J., Kusunegi, T., Oka, M., Ryu, R., Ueda, M., Matsubara, K., RIKEN; Kawai, J., Carninci, P., Adachi, J., Aizawa, K., Arakawa, T., Fukuda, S., Hara, A., Hashidume, W., Hayatsu, N., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kondo, S., Kono, H., Miyazaki, A., Oseko, N., Ota, Y., Saito, R., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T., Yoshino, M., and Hayashizaki, Y.						
TITLE	Collection, mapping, and annotation of over 28,000 cDNA clones from japonica rice						
JOURNAL	Science 301 (5631), 376-379 (2003)						
MEDLINE	22752273						

DB	Y	ORIGIN	FEATURES	COMMENT	TITLE	JOURNAL	REFERENCE	PUBMED
247	TCGTGAGACCCGCGCGCTGGAACGCGCGCCCTCGGAGCGAGGCGCTCGGAGGCAACCGG	17.5%; Score 273.2; DB 8; Length 1903; Best Local Similarity 51.2%; Pred. No. 5.2e-67; Matches 750; Conservative 0; Mismatches 698; Indels 18; Gaps 4;	1.1903 location/Qualifiers organism="Oryza sativa (japonica cultivar-group)" mol_type="mRNA" cuiilar="Nidponbora" db_xref="taxon:39947" clone="J03048K24"	URL : http://cdna01.dna.affrc.go.jp/cDNA/ NIAS Rice full-length cDNA Project Team: Kikuchi, S., Satoh, K., Nagata, T., Kawagashira, N., Doi, K., Kishimoto, N., Yazaki, J., Ishikawa, M., Yamada, H., Ooka, H., Hotta, I., Kojima, K., Shimizu, T., and Yamamoto, M. Fais Genome Sequencing & Analysis Group: Otsu, Y., Iida, Y., Fujiwara, T., Ikeda, R., Ishibiki, J., Kawamata, M., Kobayashi, M., Kodama, K., Kurokawa, T., Kusumegi, T., Lu, M., Maeda, H., Mura, J., Mizuno, K., Nakikawa, R., Nishikawa, J., Oka, M., Ryu, R., Sugano, S., Sugiyama, A., Suzuki, Y., Tsunoda, Y., Ueda, M., Xie, Q., Yokomizo, S., Yoshimura, A., Matsubara, K., and Murakami, K. Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Sequence Laboratory in Riken: Adachi, J., Aizawa, K., Akimura, T., Arikawa, T., Carninci, P., Fukuda, S., Hanagaki, T., Hara, A., Hashizume, M., Hayashida, K., Hayatsu, N., Hiramoto, K., Hirooka, T., Hori, F., Iida, J., Imanura, K., Imoto, K., Ishii, Y., Itoh, M., Kagawa, I., Kanagawa, S., Katoh, H., Kawai, T., Kishikawa-Hirozane, T., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kuribara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numasaki, R., Ohno, M., Otsu, N., Ota, Y., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Saeki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagami, T., Tagami, Y., Tagawa, A., Takahashi, F., Takuu-Akahira, S., Tanaka, T., Tomaru, A., Toyota, T., Waki, K., Yasuniishi, A., and Hayashizaki, Y.	Submitted (27-AUG-2002) Shoshi Kikuchi, National Institute of Agricultural Sciences, Department of Molecular Genetics, Head of Laboratory of Gene Expression; 2-1-2 Kamondai, Tsukuba, Ibaraki 305-8602 Japan (E-mail:ekikuchi@ias.affrc.go.jp, Tel:81-29-838-7007, Fax:81-29-838-7007) This clone is one of the 28K full-length cDNA clones from japonica rice.	Direct Submission Submitted (27-AUG-2002) Shoshi Kikuchi, National Institute of Agricultural Sciences, Department of Molecular Genetics, Head of Laboratory of Gene Expression; 2-1-2 Kamondai, Tsukuba, Ibaraki 305-8602 Japan (E-mail:ekikuchi@ias.affrc.go.jp, Tel:81-29-838-7007, Fax:81-29-838-7007) This clone is one of the 28K full-length cDNA clones from japonica rice.	128659764 2 (bases 1 to 1903) Adachi, J., Aizawa, K., Akimura, T., Arikawa, T., Carninci, P., Doi, K., Fujiwara, T., Fukuda, S., Hanagaki, T., Hara, A., Hashizume, M., Hayashida, K., Hayashizaki, Y., Hayatsu, N., Hiramoto, K., Hirooka, T., Hori, F., Hotta, I., Iida, J., Iida, Y., Ikeda, R., Imanura, K., Ishikawa, I., Imoto, K., Ishibiki, J., Ishii, Y., Ishikawa, M., Itoh, M., Kagawa, I., Kikuchi, S., Kishikawa-Hirozane, T., Kishimoto, N., Kobayashi, M., Kodama, T., Kojima, K., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kuribara, C., Kurokawa, T., Kusumegi, T., Li, C., Lu, M., Maeda, H., Matsubara, K., Matsuyama, T., Mura, J., Miyazaki, A., Mizuno, K., Murakami, K., Murata, S., Nagata, T., Nakamura, M., Nakikawa, R., Ohneda, E., Ohno, M., Ohtsuki, K., Oka, M., Ooka, H., Otsu, N., Ota, Y., Otsu, Y., Ryu, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Saeki, D., Sato, K., Satoh, K., Shibata, K., Shinagawa, A., Shiraki, T., Shishiki, T., Sogabe, Y., Sugano, S., Sugiyama, A., Suzuki, K., Suzuki, Y., Tagami, M., Tagami, T., Tagawa, A., Toyota, T., Tsunoda, Y., Ueda, M., Waki, K., Xie, Q., Yahagi, W., Yamada, H., Yamamoto, M., Yasuniishi, A., Yazaki, J., Yokomizo, S. and Yoshimura, A.	



Mon May 24 08:10:07 2004

us-09-992-901-1.rge

Page 21

```
QY 154 TATCATTTCTTCATCGGAATGTAAAGACTTGTGAAATGATGCTTAAAGCTTCTT 213
Db 307 TACGGCTCTTTCACCGCGACGTCGGGAGAAAGTCGGCTCAACCGGAGGCCCGAG 366
QY 214 CATCTTATGCTTTT---CTCTCAATATTTCTTCTAGAGTTCTCTTTTTCATC 270
Db 367 AAGCGCTGCGCTGCGCTCGACGACATATCCCGCGCTGCTCCATGTTACGAAA 426
QY 271 TGGAGAAAAATCTACGCTGCTACATTTCTGTTGGTTGGTCCAACTTTCGGTTAAG 330
Db 427 GCGCTGAGAGAGCAGCGGAAACATTCATCTGTTGGTCCCAACCGCAAGAGTATG 486
QY 331 GTAGCCGATCTCTGATTTGATCAGAGATCTCTTCTAGTCTGAGTTCTTACAGAGAAT 390
Db 487 ATTTCAGACCTGATCAATATAGGAAAGTTATGCTATATAGTTGGCCACTATGGCAAA 546
QY 391 GAAGCTACCCCTTGGTTAAACAATTGAAGGAGTGAAGTACTTACTTCTCAAGAGTAA 450
Db 547 CCAAGGCTACCCGCTCGGAAAGTTGCTAGCTCCGAGTGTAGCTATGAGAGCGAG 606
QY 451 AATGGGCTCATCTGAAAAATCATTAAGCTTCTTCTATGAGAGATCTTAAAGTTG 510
Db 607 AATGGGCAAGACCGGAGAAATTCGAAATCTGCTTTCACACGAGAAAAATTAAGCG 666
QY 511 CTGTGACCAAGTTGTTGAAGAGTGTGACTGATATGATGATTAATGGTCCGATTAAGTTA 570
Db 667 ATGCTGCAAGTTTCTTCTAATCTGCTCAAGGAAATGTTTACAGATGGGAGATTCATG 726
QY 571 TCAGAAAAAGCTGATGAGTGAATGCTATGAGTGTTCAGATTTGATGAGAT 630
Db 727 TCTATGAGAAATGTCAGAGTGAATGTTGCTGAGTTCCAAATCTTACAGAGAT 786
QY 631 GTTATGATGAGACAGCTTTTGAAGTACTGATGAAGATGTCGAGCAGTTTTCGACTT 690
Db 787 GTCATATCAAGAGACGATTCGCTACAGCTTATGAGAAAGAGAGAAATTTTTCAGCTG 846
QY 691 CAAGCTCAACAAATGCTTCTTGTGCTGAAGCTTTCAAAAAGTCTTATTCCTGCTAT 750
Db 847 CAAGCAGAGTCAAGCAGACGATTAATACAGCTTTCGACATTTTATACAGAGAT 906
QY 751 AGATTTTTCGACAAAGAGGAAATTTGAAGTCTCGAAGTGTAGCAAGAGATTAAGAA- 809
Db 907 TGTCTTCACTCACTAAAAACAGAGATTTGAGAGAAATTTGAAGAGAGCTCAGCAA 966
QY 810 -----GTGCTGTTGAAGCTGATGAGCGCGAGACAAAGCTATATGATGAGAA 861
Db 967 CTTCACAGAGAAATTAATGGAAGAGAGCGGCTATTAATAATGTGAACCACTAAT 1026
QY 862 GGGGAGAGATGTAAGAGCGCGCGAGAGATTTGTGGATTAATGATTCAAGCAAG 921
Db 1027 GGTGACTGTGTTGGCTTATTTGTGAGTCAATATAGAGAGTCAATGGGAAAGCAGA 1086
QY 922 AATG--TGACGTTTCAAGCATTTGAGAGTGAAGCTTAAAGCTTTTCTGCGCGGAAA 978
Db 1087 CTAGGATGACTCGAGCGAAATTAATGAGGATCAAGCTATTTATTTTTCAGGATG 1146
QY 979 CAGACAACTTCTATCTGCTGAGAGTGAAGACATCTTGTATATCCATGCAACCGAGATG 1038
Db 1147 GAGACAAATCAATATGCTGCTGAGATTAATTTGTCTAGATGCAACCGAGATG 1206
QY 1039 CAGGCCAAAGCAGTATAGAGTCTCAGGCTCTGCGGCTCACTGATGTCCTTACCAAG 1098
Db 1207 CAGAGCGAGCAGAGAGAGAGTCTACCACTTGG---AAGAACCAACCAAGACTAT 1263
QY 1099 GACCATGTGTTAGCTTAAAGCTTGAAGTATGATCTTGAACGAGTCTTTAAGGTTGAT 1158
Db 1264 GATAGCTTAAAGTGTGTAAGATTTGTAATGATTTCTGACGAGTTCTTAAAGTTGAT 1323
QY 1159 CACCAATTAAGTACTAGATTCAGCGCGTAAATGAGATGGAAGCTAGAGGATCAAA 1218
Db 1324 CCGCAGGTGTGTTCTTGAACAGACAGACATCAAGAAATGAGAGCTCGGCGGATCAAA 1383

QY 1219 ATCCATGTGGCAGGAGCTTCTAATCCCATATAGCGGTCCATCATGACCAAGCAT 1278
Db 1384 TATCCGCTGAGAGTACCCCTTATGCTTCCATTTTATTTATCATATATCCCATATT 1443
QY 1279 TGGGATTAATGACGTGAACGAATTCATTCAGCTCGGTTTGGGATGAGTCCCGGTCT 1338
Db 1444 TGGGAAAAAGATGAGGTGAATTCATTCAGGAGGTTGCTGATGATGCAATTCGAAGCA 1503
QY 1339 GCGAAACACCCGTTGGCTTCAATACGTTTGGCTCGAGATTGATCATGATGATGATG 1398
Db 1504 AGGAGATTCAGAGCTCTTCTTCCATTTGAGATGGGATCCCGAATCTGATCGGCGAG 1563
QY 1399 AATCTGTATATCTTCAAGCCAAATTAAGACACTGCTGATATGATTCAGCTTACCTT 1458
Db 1564 AACTTGCATTTGAGAGCCAAAGATGGCTATCTGACAACTCTTCAAGCGGTTCTCTT 1623
QY 1459 CACTTGGCTCTTATTCAGAGACAGTACAGCTTATGTTGCTTATCTCAAGAT 1518
Db 1624 GAGCTTCAACATGTAATCCAGACATTCAGAGCAATTCATGTAATCTCCAGCAGAGAT 1683
QY 1519 GGTGACCAATCACCTTCCGAGATT 1544
Db 1684 GGTGCAAAATTAAGCTGAAGAAAT 1709

Search completed: May 22, 2004, 19:00:48
Job time : 6292 secs
```

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 22, 2004, 10:50:51 ; Search time 681 Seconds  
(without alignments)  
9750.270 Million cell updates/sec

Title: US-09-992-901-1  
Perfect score: 1563  
Sequence: 1 atgaggaagaagatgacag.....tgaccatcatgagatcga 1563

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 212409041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : N\_Geneseq\_25Jan04:\*

- 1: geneeqn1980s:\*
- 2: geneeqn1990s:\*
- 3: geneeqn2000s:\*
- 4: geneeqn2001s:\*
- 5: geneeqn2002s:\*
- 6: geneeqn2003s:\*
- 7: geneeqn2003bs:\*
- 8: geneeqn2003cs:\*
- 9: geneeqn2003cs:\*
- 10: geneeqn2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match Length	DB ID	Description
1	1563	100.0	1563 3 AAA99072	AAA99072 Arabidops
2	318.8	20.4	1539 6 ABK50985	ABK50985 DNA encod
3	270.4	17.3	1584 6 ABL60758	ABL60758 Malze cyt
4	269.2	17.2	1539 6 ABK50986	ABK50986 DNA encod
5	265.4	17.0	1833 6 ABK50983	ABK50983 Zea mays
6	241.2	15.4	1557 6 ABZ12287	ABZ12287 Arabidops
7	241.2	15.4	1712 6 AAC43309	AAC43309 Arabidops
8	206.2	13.2	1365 8 ADA48311	ADA48311 Rice gene
9	194.8	12.5	1539 6 ABZ12574	ABZ12574 Arabidops
10	185.4	11.9	1695 3 AAC41990	AAC41990 Arabidops
11	184	11.8	588 3 AAA99071	AAA99071 Arabidops
12	153.6	9.1	1124 6 ABN98320	ABN98320 Arabidops
13	141.8	9.1	1377 8 ADA48675	ADA48675 Rice gene
14	104.2	6.7	1456 3 AAC54724	AAC54724 Arabidops
15	102.6	6.6	4458 3 AAC53115	AAC53115 Arabidops
16	100.6	6.4	333 6 ABK50982	ABK50982 Zea mays
17	93	6.0	436 7 ABK61908	ABK61908 Arabidops
18	88.6	5.7	474 6 ABK61933	ABK61933 Arabidops
19	88.6	5.7	475 7 ABK60704	ABK60704 Arabidops
20	77.4	5.0	435 7 ABK61927	ABK61927 Arabidops
21	75.8	4.8	612 3 AAA68008	AAA68008 Eucalyptu
22	75.8	4.8	612 3 ADD41758	ADD41758 Comnarte
23	73.4	4.7	398 3 AAA68076	AAA68076 Eucalyptu

24	73.4	4.7	398 9 ADD41826	ADD41826 Ferulate-
25	72.2	4.6	422 7 ABK62213	ABK62213 Arabidops
26	72.2	4.6	424 7 ABK61377	ABK61377 Arabidops
27	70.6	4.5	1392 4 AAS02937	AAS02937 Heme doma
28	70.6	4.5	3150 4 AAF54832	AAF54832 Nucleotid
29	70.6	4.5	3150 4 AAF54832	AAF54832 B. megate
30	70.6	4.5	3150 5 AAF26324	AAF26324 B. megate
31	70.6	4.5	3225 7 ABZ70676	ABZ70676 Bacillus
32	70.6	4.5	4957 7 AAL51096	AAL51096 Bacillus
33	70.6	4.5	4957 7 AAL54548	AAL54548 DNA of P4
34	69	4.4	3150 3 AAA46648	AAA46648 Nucleotid
35	69	4.4	3150 4 AAD02365	AAD02365 Bacillus
36	69	4.4	3150 7 ACC48000	ACC48000 B. megata
37	68.6	4.4	285 7 ABK84315	ABK84315 Corn ear-
38	63	4.0	1936 9 ADB59019	ADB59019 Toxicity
39	63	4.0	1936 9 ADB53777	ADB53777 Primary r
40	60.6	3.9	180 6 ABN75725	ABN75725 Human cyt
41	57.2	3.7	1494 6 ABZ13577	ABZ13577 Arabidops
42	57.2	3.7	1494 6 ADA68047	ADA68047 Arabidops
43	55.2	3.5	497 8 ADA49004	ADA49004 Wheat gen
44	53.8	3.4	1976 6 ABK48198	ABK48198 Human CDN
45	53.4	3.4	631 6 ABQ66106	ABQ66106 Arabidops

## ALIGNMENTS

RESULT 1  
ID AAA99072 standard; DNA; 1563 BP.  
XX AAA99072;  
XX 18-JAN-2001 (first entry)  
XX Arabidopsis thaliana basl nucleotide sequence SEQ ID NO.1.  
XX Arabidopsis thaliana; basl; promoter; cytochrome P450; CYP72B1; plant;  
XX brassinosteroid signalling; brassinosteroid synthesis; brassinolide; ds.  
XX Arabidopsis thaliana.  
XX OS  
XX WO200055302-A2.  
XX PD 21-SEP-2000.  
XX PF 16-MAR-2000; 2000WO-US006915.  
XX PR 16-MAR-1999; 99US-0124570P.  
XX PR 14-DEC-1999; 99US-0170931P.  
XX PR 20-DEC-1999; 99US-0172832P.  
XX PA (SALK ) SALK INST BIOLOGICAL STUDIES.  
XX Neff MM, Chory J;  
XX WPI; 2000-638195/61.  
XX P-PDB; AAB23917.  
XX Transgenic plants having modulated brassinolide synthesis resulting in  
XX insect resistance, dwarfism and darker-green foliage compared with wild-  
XX type plants, have nucleic acid encoding Basl polypeptide in its genome.  
XX Disclosure; Fig 1A; 104pp; English.  
XX The present invention describes a genetically modified plant (1)  
XX comprising at least one exogenous nucleic acid sequence encoding a Basl  
XX polypeptide, homologous or functional fragment, in its genome or at least  
XX a regulatory sequence that modulated expression of endogenous basl gene,  
XX homologous or functional fragment, and which is characterised as having  
XX modulated brassinolide activity or synthesis. The basl gene encodes a  
XX cytochrome P450 (CYP72B1), which has a role in brassinosteroid signalling  
XX or synthesis. Overexpression of the basl gene in plants causes a dark

CC green, dwarf phenotype which mimics plants that have low levels of the  
CC plant hormone, brassinolide. Overexpression of the *bast1* gene also  
CC increases resistance to insects in plants. The present sequence  
CC represents the *Arabidopsis* *bast1* nucleotide sequence, which is used in the  
CC exemplification of the present invention

Sequence 1563 BP; 424 A; 310 C; 377 G; 452 T; 0 U; 0 Other;

Query Match	100.0%	Score 1563;	DB 3;	Length 1563;
Best Local Similarity	100.0%	Pred. No. 0;		
Matches 1563;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

QY	ATGAGAGAAAGAAAGTAGCAGCTGCTTCAATCCAAAGGCTCTGGTCTGATCAATCTTA	60
Db	1 ATGAGAGAAAGAAAGTAGCAGCTGCTTCAATCCAAAGGCTCTGGTCTGATCAATCTTA	60
QY	AGTCTTGTAATAGTGAAGGGATGCTCTGTTATGTTGAGACCAAGAAAGATTGAAGA	120
Db	61 AGCTCTGTATAGTGAAGGGATGCTCTGTTATGTTGAGACCAAGAAAGATTGAAGA	120
QY	CATTCTCTAAACAAAGGAATTCGAGGCTCCTTATCATTTCTTCATCGGAATGTTAA	180
Db	121 CATTCTCTAAACAAAGGAATTCGAGGCTCCTTATCATTTCTTCATCGGAATGTTAA	180
QY	GAACTTGATGATGATGCTTAAAGCTTCTTCATCCATCCATGCTTCTCCAAATAT	240
Db	181 GAACTTGATGATGATGCTTAAAGCTTCTTCATCCATCCATGCTTCTCCAAATAT	240
QY	CTTCCCTAAGCTCTCTCTTTTAAACAATACGAGAGAAAATCTACGCTGTAATCTTG	300
Db	241 CTTCCCTAAGCTCTCTCTTTTAAACAATACGAGAGAAAATCTACGCTGTAATCTTG	300
QY	GTTGGTTCGCTCCAACTTCCGTTTAAACGTAAGCCATCTCTGATTTGATCAGAGATC	360
Db	301 GTTGGTTCGCTCCAACTTCCGTTTAAACGTAAGCCATCTCTGATTTGATCAGAGATC	360
QY	TTCTCTAAGCTCTGAGTTCTTACAGAAAGATGAGCTCACTCTTGTTAAACAATGAA	420
Db	361 TTCTCTAAGCTCTGAGTTCTTACAGAAAGATGAGCTCACTCTTGTTAAACAATGAA	420
QY	GCGATGAGTACTTGTAGCTCAAAAGGTGAAAAATGGGCTCATCTCAAAATCATTAAC	480
Db	421 GCGATGAGTACTTGTAGCTCAAAAGGTGAAAAATGGGCTCATCTCAAAATCATTAAG	480
QY	CCTACTTTCAATGAGAAATCTTAAAGTCTGTACAGTTGCTTGAAAGATGAGACT	540
Db	481 CCTACTTTCAATGAGAAATCTTAAAGTCTGTACAGTTGCTTGAAAGATGAGACT	540
QY	GATAGGTGATTAAGTGTCCGATTAAGTATCAGAAAACGCTGAAGTTGAGTAGATGC	600
Db	541 GATAGGTGATTAAGTGTCCGATTAAGTATCAGAAAACGCTGAAGTTGAGTAGATGC	600
QY	TATGAGTGTTCAGATTTTGACTGAAGATGTTATTAGTAAACAGCTTTTGAAGTAGC	660
Db	601 TATGAGTGTTCAGATTTTGACTGAAGATGTTATTAGTAAACAGCTTTTGAAGTAGC	660
QY	TATGAGATGCTGAGAGAGTTTTTGCACCTTCAAGCTCAACAATAGCTCTTGAGCTGAA	720
Db	661 TATGAGATGCTGAGAGAGTTTTTGCACCTTCAAGCTCAACAATAGCTCTTGAGCTGAA	720
QY	GCTTTGAAAAAGTCTTCATCTCGCTATAGATTTTTCCGACAAAGGGAAATTTGAG	780
Db	721 GCTTTGAAAAAGTCTTCATCTCGCTATAGATTTTTCCGACAAAGGGAAATTTGAG	780
QY	TCTTGGAAGTTTAAACAAGAGATTAAGAAAGTGGTTGGAAGCTGATGAGCGCGGAGA	840
Db	781 TCTTGGAAGTTTAAACAAGAGATTAAGAAAGTGGTTGGAAGCTGATGAGCGCGGAGA	840
QY	CAAAACCTTATAGTGAAGAGGAGAAATGTAAAGACCGGCGCGAAGATTTGTTG	900
Db	841 CAAAACCTTATAGTGAAGAGGAGAAATGTAAAGACCGGCGCGAAGATTTGTTG	900
QY	GGATTTATGATTCAGCGAAAGAAATGTGACCGTTACAGACATTTGTGAGAGATGTAAAGC	960

[illegible]

XX	07-DEC-2000; 2000US-0251518P.
XX	(CANADA) NAT RES COUNCIL CANADA.
XX	Krochko JE, Cutler AJ, Abrams SR;
XX	WPI; 2002-519663/55.
DR	P-PsDB; AAU97096.
XX	New isolated and purified DNA that encodes protein having abscisic acid
PT	(ABA) hydroxylase activity, useful for altering catabolism of abscisic
XX	acid in plants.
XX	Claim 7; Page 104-105; 117pp; English.
CC	The invention relates to an isolated and purified DNA (I) that encodes a
CC	protein having abscisic acid (ABA) hydroxylase activity. (I) is useful
CC	for producing a transgenic plant which involves introducing (I) into a
CC	genome of the plant or its part, and carrying out plant growth and
CC	development. (I) is useful for modifying catabolism of ABA or ABA
CC	analogues in a plant. ABA hydroxylase clone pBE10-30-3 cDNA sequence is
CC	useful for identification of related sequences from other plant species.
CC	(I) is also useful for altering ABA and ABA analog catabolism in plants.
CC	The present sequence encodes abscisic acid (ABA) hydroxylase CYP72A7
SQ	Sequence 1539 BP; 452 A; 307 C; 371 G; 409 T; 0 U; 0 Other;
Query Match	20.4%; Score 318.8; DB 6; Length 1539;
Best Local Similarity	52.8%; Pred. No. 7.1e-85;
Matches 761; Conservative	0; Mismatches 672; Indels 9; Gaps 3
Dy	90 GTTATGAGTAGACCAAGAAGATTGAAGAACTTTCTTAACAAGATTGAGTGCC 149
Qy	78 GCCTCGATTAAACCAAAGATCTTGAGAGTCTTTGAAAAGACGGGTCTTCACGGAAC 137
Dy	150 TCCTTATCATTTCTTCATGGAATGTTAAAAACTGTTGGAATGATGCTTTAAAGCTTC 209
Qy	138 TCCCTACACCTCCTCTGCTGGAGATATAAGAAAGATTGATATGATGATGGAACGAG 197
Dy	210 TTCTCATCCGATGCCCTTCTCTCACAAATTTCTCTAGAGTTCTCTCTTTTACCANTCA 269
Qy	198 ATCTTAACCCATCAATGTAACGAGATGATCACCCACGCTCTCTCTGCTTAA 257
Dy	270 CTGAGAAAAATCTACGAGTCAATTTCTGGTTTTGTTGGTGCCTCACTTCCGGTTAAC 329
Qy	258 GATGCTCAATTTCTACGGAAGACCTTTCTATGATGATGAGACATTCCTCCAGATTGT 317
Dy	330 GGTAGCGCATCCGATTTGATACAGAGATCTCTCTAAGCTGAGTCTTACAGAAAGAA 389
Qy	318 GATTAAGCAATCTGACAGATATAGGAAGTCTTTATTAAGTACAGACCTTTGAGAAAC 377
Dy	390 TGAAGTCAACCTTTGGTTTAAACAATTGAAGCGATGGACTATTAGTCTCAAGGTGA 449
Qy	378 TTCTACATTCCTTTGATCAGATGTTGACAGS--TGCGCTTGCAAGTTACAGAGAGA 434
Dy	450 AAAATGGCTCATATCGAAAAATCATTAAGCCCTAATTTTCAATGAGAAATCTTAAGTT 509
Qy	435 TAATGGCGCAGTCAAGAGATATCAACCCGGCTTTTCACTCGAAAAATCAAGAA 494
Dy	510 GCTTGTACACAGTTGTGTTAAGAGTGTACTGATATGGTGATTAATGATCCGATAAGTT 569
Qy	495 CATGATCCCGGCTTTACCAATTTGTCAGGAGAGTTGTCTCAATGGAGAAAGTATT 554
Dy	570 ATCAGAAAAG--GTGAAGTTGAGTAGATGTCTAGAGTGTTCAGATTTTGAATGA 626
Qy	555 TAAAGATTAAGAAATCCCTCTTGAAGTGAATGTTGGCTTGGCTTGAATGAATGCG 614
Dy	627 AGATGTTATTAGTAAGAACAGCTTTTGAAGTAGTATGATGATGATGTCAGACGTTTTCG 686
Qy	615 GGATGTCATCTCACAATCTGTTTGAAGTAGTATTAAGAAAGGACAGAAATATTCCA 674
Dy	687 ACTTCAAGTCAACAAATGCTTTTGTGCTGAGACTTTCAAAAAGTCTTCAATCTCG 746

Db	675	ACTACAAAGGGAAATTGGCTGAGCTTATCGACAAGCTTTTAAGAAATCTTACATCCCTGG	734
Qy	747	CTATGATTTTTTCCGACAAAGAGSAAATTTGAAGTCTCGSAGATTGACAAGAGATTAAG	806
Db	735	ATCGAGGTTTTACCCAAACAAAGGCAATAGAAAGATGAAGCAATAGATTGAGAATAGA	794
Qy	807	GAAAGCTGTTGTTAAAGCTGATAGAGCGGCGGAGACAAAACGTTATAGATGAGNAAAGGGA	866
Db	795	CGTAATATTGAGAGGTATTGTGACAAACCGGAGAAAGGAGAGAAAGCTGGAGAAACGAC	854
Qy	867	AGAAATGTAGAGACCGCGGCGGAGAAAGATTGTTGGGATTATAGA--TTCAAGCCAAAGAA	923
Db	855	AAACCATGATTTTGTGGGGAATTTGCTGGAATCCATTTCAAGAGAAATCTCAAGAAATACG	914
Qy	924	TGTGACGTTGAGACATTTGTGAGAGAGTGTAAACCTTTTCTTTCGCCGGAACACAC	983
Db	915	AATGAGCGTAGAAGATGTGATGAAGAGTGCAAAGTTGTTTATTTCGGGGACAAAGAAC	974
Qy	984	AACTCTAAATCGCGTGAAGTGGACGACATCTTGATCCATGACACCCGSAATGAGCAGC	1043
Db	975	AACTTCAGTACTTTTGTGTGAGCATAGTATTTTATTTAACCATACACAAAGCTGGCAAGC	1034
Qy	1044	CAAGACACGTGATGAGGTCTCAGAGGTCTGGGCTCAAGTATGTCCTTACCAAGACCA	1103
Db	1035	TCGCGCACGAGAGGAAGTGAATGCAGATGACTCGGTGAAATTAATTAACGTATGTGACTC	1094
Qy	1104	TGTGCTAAGCTTAAACGTTGAGATGATCTTGAAAGAGCTCTTAAAGGTGTATCCACC	1163
Db	1095	CCTTAACACCTTAAGTCATGACTATGATCTTCAATGAGATTGAGGCTTAATACCTCC	1154
Qy	1164	AATAGTAGCTACGATTCGACGCGCTAATCGATGTAAGTGAAGCTTAGAGGGTTACAAATCC	1223
Db	1155	GGTGGCTCAGCTTAAACGAGTGTCCAAAGAAAGAAAGCTCGGAGAGTTGACCCCTTCC	1214
Qy	1224	ATGTGCGACGAGAGCTTCAATCCCAATCATAGAGGTCATCATGACCAAGCATTGGGCG	1283
Db	1215	AGCTGGAATTCAAATTTATCTTACCAACTATTTCTTGTCAGCGTGACACCGAGCTTTGGGG	1274
Qy	1284	TATGACGTGACGAAATTCATTCAGCTCGGTTTGGGATGAGAGTGCGCGGTGCTGCCAA	1343
Db	1275	CGATGATGAGCGGATTTTAACCCGAGCGGTTCAAGAGACGGGCTCTCAAAAGCAACAA	1334
Qy	1344	ACACCCCGTGTGCTTCAACCGTTTGGCCTCGAGAGTGTGAATGATGATGGTCAGAAATCT	1403
Db	1335	GAAACGAGCTCTTCTTCTTCTCCCTTGGATGGGAACTTAGAATTTGCATGGGTCAAGATT	1394
Qy	1404	TGCTATACTTCAGGCCAATTTGACACTCGCTGTATGATCCAAACGCTTCACTTCACTT	1463
Db	1395	TGCTATGTTGAGGCAAAAGATGGCAATGGCTTTGATTCAAAAAGTTCTCTCTTGAACT	1454
Qy	1464	GAGCTCTACTTATACGACATGACCTTACCGCTCTTATGTTGCTTTATCCCAACATGATGTC	1523
Db	1455	CTCTCCCTTCTTATGTTTCAGCGCTCAAAACATGATGACCACTGCTCCCAATTTGGAGGC	1514
Qy	1524	AC 1525	
Db	1515	TC 1516	
RESULT 3			
ABL60758			
ID ABL60758 standard; cDNA; 1584 BP.			
XX ABL60758;			
XX AC			
XX DT 10-SEP-2002 (first entry)			
XX DE Maize cytochrome P450, CYP72A1 coding sequence.			
XX KW Maize; cytochrome P450; CYP72A1; CYP92A1; plant; gene therapy; anticdote;			
XX gene; ss.			
XX			

Mon May 24 08:10:07 2004

us-09-992-901-1.rng

Page 4

OS Zea mays.  
XX Key 1.1584 Location/Qualifiers  
FH CDS 1.1584  
FT /tag=a  
FT /product="CYP72A1"  
XX US6380465-B1.  
XX 30-APR-2002.  
XX 11-JUL-1999; 99US-00351229.  
XX 12-JUL-1998; 98US-0092596P.  
XX (KENT) UNIV KENTUCKY RES FOUND.  
XX Barrett M;  
XX MPI: 2002-470227/50.  
XX P-PSDB; ABB08079.  
XX  
XX Nucleic acid molecules encoding a Zea mays cytochrome P450 CYP72A1,  
XX useful for the production of plants resistant to heterocyclic,  
XX sulfonylurea and substituted urea herbicides and organophosphate  
XX insecticides e.g. Classic and Pursult.  
XX  
XX Claim 11; Col 9-12; 21pp; English.  
XX  
XX The invention relates to isolated nucleic acid molecules encoding Zea  
XX mays (maize) cytochrome P450 designated CYP72A1 and CYP92A1. The  
XX polypeptides can be expressed by standard recombinant methodology. The  
XX nucleic acids may be used in genetic engineering protocols to transform  
XX plants and other eukaryotes e.g. Yeast, maize (especially), soybean,  
XX beer, tobacco, wheat, barley, poppy, rape, sunflower, alfalfa, sorghum,  
XX rose, carnation, gerbera, carrot, tomato, lettuce, chicory, pepper, melon  
XX and cabbage. They may be used in this way to confer protection against  
XX heterocyclic herbicides, sulfonylurea herbicides, substituted urea  
XX herbicides and/or organophosphate insecticides. The present sequence  
XX represents the Z. mays CYP72A1 coding sequence  
XX  
SQ Sequence 1584 BP; 430 A; 389 C; 414 G; 351 T; 0 U; 0 Other;  
Query Match 17.3%; Score 270.4; DB 6; Length 1584;  
Best Local Similarity 51.5%; Pred. No. 2.7e-70;  
Matches 680; Conservative 0; Mismatches 626; Indels 15; Gaps 2;  
QY 228 CTCACACATATCTCTCTAGAGTCTCTCTTTTACCATCACTGAGAAATCTACGG 287  
DB 249 CTGCACGACATCACCCACGCGTGCAGCCCATCATCAGACACCATCAAGAAATACGG 308  
QY 288 TGCTACATTTCCGTTGGTTCGGTCCACCTTCCGGTTAAGCGTACCGATCCGATT 347  
DB 309 GAACATATCGTTCACCTGGTTCGCGCCACACCAAGGATATATCTCTGACCCAGATT 368  
QY 348 GATCAGAGAGATCTTCTCTAAGTGTGAGTTCTACGAGAGATGAAGTCACTACCCCTTGGT 407  
DB 369 AGTCAAGAGGCTGCTCTATATAGTTGGCCACTTGGCAACCAAGAGATGACCGGAT 428  
QY 408 TAAACCACTTGAAGCGGATGAGTACTAGTCTCAAGGTAAGGAAATGGGCTCATCATCG 467  
DB 429 TGGAGGTTGCTAGCGAACGGGCTTGAATCATGATGTAAGGAAATGGGCAACGACG 488  
QY 468 AAAAATCATTACCCCTACTTTTCAATAGAGAAATCTTAAGTGTCTGTAACGATGGT 527  
DB 489 GAGATCTCTATCTGATTCACATGAAAAAATGAAGGATATGCCGATATTTTC 548  
QY 528 GAAGAGTGTGACTATAGTGTGATTAATGATGTCGCAATGATTACAGAAACGATGAAGT 587  
DB 549 TACCGTGTGATTAAGATTAATGATTAATGATTAATGATTAATGATTAATGATTAATG 608  
QY 588 TGAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 647

DB 609 TGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 668  
QY 648 TTTTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 707  
DB 669 GTTTGGAGCACTATCAAGAGGAGAGATTTTGGATGATCAAGAGATGATGATGATGATG 728  
QY 708 TCTTGTGCTGAGCTTTTCAAAAAGTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 767  
DB 729 AGCCCTCATCCATCTGTTTCAAGCAATTTATCCAGGCTATGTTCTTCCGCCAAC 788  
QY 768 AGGAAATTTGAAGTCTCGAAAGTTAGCAAGAGATTAAGAAAGTCTGTTGAAGCTGAT 827  
DB 789 AAACACAGAGAGATGAGAGCAATGATGATGATGATGATGATGATGATGATGATGATG 848  
QY 828 AGAGCGCGGACACAAACGCTATATAGTGAAGAGGGAAGATGTAAGAG----- 879  
DB 849 TGGAAAGAGAGAGAGATGATCAAAAACAGAGAAACAAATAGATGATGATGATGATG 908  
QY 880 -----CGGCGCGGAGAGATTTTGGATTAATGATTAATGATTAATGATTAATGATTA 935  
DB 909 ATTACTGAGTAAACACAGGCAATCAATGCAATGCAAGCTGGGATTTGCAACAG 968  
QY 936 GACATTTGTGAGAGATGTAAGCTTTTCTTCCGCGGAAACAGCAACTTCTATCT 995  
DB 969 AGATGATGATGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1028  
QY 996 GCTGACGTGAGAGCACTTGTGATGATGATGATGATGATGATGATGATGATGATGATG 1055  
DB 1029 GCTTACTGAGACTTATTTGTTGATGATGATGATGATGATGATGATGATGATGATG 1088  
QY 1056 TGAAGTCTCAGAGTCTGCGGCTCAAGTATGCTTCAAGAGACCATGCTTAAGCT 1115  
DB 1089 AAGGTTTGGACCTTTTG-----AAGAACACACAGATTAATGATGATGATGATG 1145  
QY 1116 TAAACGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1175  
DB 1146 CAGAGTGTACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1205  
QY 1176 GATTGAGCGGTAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1235  
DB 1206 AACCAAGAGATCTTAAAGAAATGAGAGCTCGGATGATCAATATCTGAGAGATG 1265  
QY 1236 GCTTCAATCCCAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1295  
DB 1266 GCTTCTTCTGCGGCTATCTTCACTCACTGATCCGATTTGGGAGAAAGACCAAG 1325  
QY 1296 GCAATTCATCCAGCTCGTTTGGGATGAGTGCAGCTGCTGCCAACAACCCGTTGG 1355  
DB 1326 CGAGTTCAACCCAGAAAGTTTGGCAACGGCATCTCCAGGCAACAGGATCAGGCTC 1385  
QY 1356 CTTCAATCCGTTTGGCTCGAGTTGCTATGATGATGATGATGATGATGATGATGATG 1415  
DB 1386 TTTCTTCCGTTCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1445  
QY 1416 GGCCAATTTGACCTCGCTGATGATGATGATGATGATGATGATGATGATGATGATG 1475  
DB 1446 AGCCAAAGATGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1505  
QY 1476 TCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1535  
DB 1506 CAGCCAGCGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1565  
QY 1536 C 1536  
DB 1566 C 1566  
RESULT 4  
ABK50986  
ID ABK50986 standard; DNA; 1539 BP.  
XX AC ABK50986;  
XX

DT 24-SEP-2002 (first entry)  
 XX DNA encoding abscisic acid (ABA) hydroxylase CYP72A14.  
 DE ABA hydroxylase; abscisic acid hydroxylase; transgenic; plant; CYP72A14;  
 KW Gene; ds.  
 XX Arabidopsis thaliana.  
 OS  
 PH Key Location/Qualifiers  
 FT CDS 1..1539  
 FT /tag= a  
 FT /product= "ABA hydroxylase, CYP72A14"  
 XX  
 XX W0200246377-A2.  
 XX  
 XX 13-JUN-2002.  
 XX  
 XX 06-DEC-2001; 2001MO-CA001756.  
 XX  
 XX 07-DEC-2000; 2000US-0251518P.  
 XX  
 XX (CAN ) NAT RES COUNCIL CANADA.  
 XX  
 XX Krochko JE, Cutler AJ, Abrams SR;  
 XX  
 XX WPI; 2002-519663/55.  
 DR P-PSDB; AAU97097.  
 XX  
 PT New isolated and purified DNA that encodes protein having abscisic acid  
 PT (ABA) hydroxylase activity, useful for altering catabolism of abscisic  
 PT acid in plants.  
 XX  
 XX Claim 13; Page 104-105; 117pp; English.  
 XX  
 CC The invention relates to an isolated and purified DNA (I) that encodes a  
 CC protein having abscisic acid (ABA) hydroxylase activity. (I) is useful  
 CC for producing a transgenic plant which involves introducing (I) into a  
 CC genome of the plant or its part, and carrying out plant growth and  
 CC development. (I) is useful for modifying catabolism of ABA or ABA  
 CC analogues in a plant. ABA hydroxylase clone pBE10-30-3 cDNA sequence is  
 CC useful for identification of related sequences from other plant species.  
 CC (I) is also useful for altering ABA and ABA analog catabolism in plants.  
 CC The present sequence encodes abscisic acid (ABA) hydroxylase CYP72A14  
 XX  
 XX Sequence 1539 BP; 452 A; 330 C; 369 G; 388 T; 0 U; 0 Other;  
 SQ  
 Query Match 17.2%; Score 269.2; DB 6; Length 1539;  
 Best Local Similarity 51.1%; Pred. No. 6.2e-70;  
 Matches 737; Conservative 0; Mismatches 693; Indels 12; Gaps 4;  
 QY 90 GTATAGTGGAGACCAAGAAAGATTGAAGAACTTCTTAACAGGAATTCGAGTCC 149  
 DB 81 GGTGTTGTTCAACCAAGATGCTTGAGCGTCCCTGAGAGCAAGCTTTCCGAAAC 140  
 QY 150 TCTTATCATTTTCTCATCGGAATGTTAAGAACTTGTGAGTATGCTTAAAGCTTC 209  
 DB 141 TCTTACAGCGCTTAATCCGCGATTTTAAAGATGATCAGATGTTCAATGAGCAAC 200  
 QY 210 TTTTCATCTTATGCTTTCTCTCAATATTTCTTCAAGTCTCTCTTTTACCATCA 269  
 DB 201 ATCCAAACCCATCAACCAACAGATGATCAACCCCTGATAGCTCATCCCTTGA 260  
 QY 270 CTGAGAAATTTCAAGGTGCTCATTTCTGTTGGTTCGTCACACTTCCGTTAAC 329  
 DB 261 AATGCTCAAGCTCAAGAGAGATTAATTAAGTGTGAGCAATACCAATATCAG 320  
 QY 330 GGTAGCCGATCTGATTTGATCAGAGAGATTTCTTAAGTCTGAAGTTCTAGAGAGAA 389  
 DB 321 CATATGATCTCTGCAATATCAAGAAAGTTCACAAAGTATGATCTTCCAGAGGC 380  
 QY 390 TGAAGCTACCTTGTGTTAAACAATTGAAGCGATGAGTACTTATGCTCAAGAGTA 449

DB 381 GCATACGTTTCTTTACCAAAATTAATCT---AGGCACGGAGACTCGTTAGTATGATGGCGA 437  
 QY 450 AAAATGGGCTCATCAACCAAAATTAATAGCTTATTTTCAATGAGAAATCTTAAATT 509  
 DB 438 TAAATGGGCGCAACCCGAAATATCAATCCGCTTTTCCACCTTGAGAGATCAAGAA 497  
 QY 510 GCTTGATCAAGTGTGTTGAAGAGTCACTATATGTTGAGTAAATGTTCCGATTAATT 569  
 DB 498 TATGTAATGTTTTCACGAAGAGTCAAGAGCTGAGTGTGTTGATGAGGCAAGTACT 557  
 QY 570 ATCAGAAACGGTGA---AGTTGAGTATGATGATGATGATGATGATGATGATGATG 626  
 DB 558 CTGGATTAAGAGGCTCATGATGATGATGATGATGATGATGATGATGATGATGATG 617  
 QY 627 AGATGTTATTAATGAAACAGCTTTTGAAGTATGATGATGATGATGATGATGATGAT 686  
 DB 618 AGATGATCTCCGATCTGCTTTTGTGATGATGATGATGATGATGATGATGATGATG 677  
 QY 687 ACTTCAAGCTCAAAATGCTTCTTGTGATGATGATGATGATGATGATGATGATGATG 746  
 DB 678 ACTCAGGAGAACTTATGACAGCTGATGATGATGATGATGATGATGATGATGATG 737  
 QY 747 CTATAGATTTTTCGACAGAGGAAATTTGAAGTCTCGAAGTTATGACAGAGATTAAG 806  
 DB 738 ATATATTATCTCCCAACAAAGGTTAATGAGATGAAACAGCAGCCAGAGAAATCCA 797  
 QY 807 GAAGTGTGTTGAAAGCTGATGATGATGATGATGATGATGATGATGATGATGATG 866  
 DB 798 AGATATATGAGAGGATATATTAACAAAGGAGAGAGAGAGAGAGAGAGAGAGAG 857  
 QY 867 AGATATGAG 925  
 DB 858 AAGCAGAGATTTGCTAGTATCTTCTGAATCAAACTTGGGCAAGAGAGAGATGG 917  
 QY 926 --TGACGTTTCAGACATGTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 983  
 DB 918 AATGATATCCGAGATATATGATGATGATGATGATGATGATGATGATGATGATG 977  
 QY 984 AACTTAAATCTGCTAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1043  
 DB 978 AATATAGATCTTCTGTTTGAAGATGATGATGATGATGATGATGATGATGATG 1037  
 QY 1044 CAAAGCAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1103  
 DB 1038 TCGTGACAG 1097  
 QY 1104 TGTGTTAAGCTTAAACGTTGATGATGATGATGATGATGATGATGATGATGATG 1163  
 DB 1098 TAAAC---CAACTAAAGTTATGATGATGATGATGATGATGATGATGATGATG 1154  
 QY 1164 AATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1223  
 DB 1155 TGTATGCTCAAGCTACCCGAGCAATTCACAAAGAGAGAGAGAGAGAGAGAGAG 1214  
 QY 1224 ATGTGACAGAGAGCTTCAATCCATCAATGAGAGAGAGAGAGAGAGAGAGAGAG 1283  
 DB 1215 AGCGAGTGTCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 1274  
 QY 1284 TAAATGAGTAAAGAGATTAATCAATCAATCAATCAATCAATCAATCAATCAAT 1343  
 DB 1275 AAAGATGAG 1334  
 QY 1344 AAGACCGGTTGCTTCAACAGGTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1403  
 DB 1335 GAACCAAGTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1394  
 QY 1404 TGCTATATCTTCAAGCCAAATTAATCAATCAATCAATCAATCAATCAATCAAT 1463  
 DB 1395 TAAATGCTTGAAGCAAAATGAGCAATGATGATGATGATGATGATGATGATGAT 1454  
 QY 1464 GGCCTCTATCTTATCAAGATCAATCAATCAATCAATCAATCAATCAATCAAT 1523  
 DB 1455 TTCTCTTCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1514

QY 1524 AC 1525  
 DB 1515 TC 1516

RESULT 5  
 ABK50983  
 ID ABK50983 standard; cDNA; 1833 BP.  
 XX ABK50983;  
 XX  
 XX 24-SEP-2002 (first entry)  
 DE Zea mays abscisic acid (ABA) hydroxylase cDNA clone pBE10-30-3.  
 XX  
 XX ABA hydroxylase; abscisic acid hydroxylase; transgenic; plant;  
 KM pBE10-30-3; corn; gene; ss.  
 XX  
 XX Zea mays.  
 OS  
 XX MO200246377-A2.  
 XX  
 XX 13-JUN-2002.  
 PD  
 XX 06-DEC-2001; 2001WO-CA001756.  
 PF  
 XX 07-DEC-2000; 2000US-0251518P.  
 PR  
 XX (CANA) NAT RES COUNCIL CANADA.  
 PA  
 XX Krocenko JE, Cutler AJ, Abrams SR;  
 PI WPI; 2002-519663/55.  
 XX  
 XX New isolated and purified DNA that encodes protein having abscisic acid  
 PT (ABA) hydroxylase activity, useful for altering catabolism of abscisic  
 PT acid in plants.  
 XX  
 XX Claim 1; Page 103-104; 117pp; English.  
 PS  
 XX  
 CC The invention relates to an isolated and purified DNA (I) that encodes a  
 CC protein having abscisic acid (ABA) hydroxylase activity. (I) is useful  
 CC for producing a transgenic plant which involves introducing (I) into a  
 CC genome of the plant or its part, and carrying out plant growth and  
 CC development. (I) is useful for modifying catabolism of ABA or ABA  
 CC analogues in a plant. ABA hydroxylase clone pBE10-30-3 cDNA sequence is  
 CC useful for identification of related sequences from other plant species.  
 CC (I) is also useful for altering ABA and ABA analog catabolism in plants.  
 CC The present sequence represents abscisic acid (ABA) hydroxylase cDNA  
 CC clone pBE10-30-3.  
 CC  
 SQ Sequence 1833 BP; 497 A; 436 C; 459 G; 441 T; 0 U; 0 Other;  
 Query Match 17.0%; Score 265.4; DB 6; Length 1833;  
 Best Local Similarity 51.3%; Pred. No. 9.5e-69;  
 Matches 676; Conservative 0; Mismatches 626; Indels 15; Gaps 2;

QY 232 CACATATTTCTTCAGAGTTCTCTTTTACATCACTGAGAAAAATCTACGGTCT 291  
 DB 303 CACGACATCGCCCGCGGTGCAAGCCATGATCAAGACCACTCAAGGAATTCGGAAA 362  
 QY 292 AATTTCTGTTGGTTGGTTCGTCACATTTCCGGTTAACGGTTACCGATCTGATTGATC 351  
 DB 363 CTATCGTTCACCTGGTTGGCCAAACCAAGGGTATGATTCCTGACCAAGATTAGTC 422  
 QY 352 AAGAGATCTTCTTAAGTCTGAGTTCTACGAGAAAGATGAAGCTACACCTTGGTTAA 411  
 DB 423 AAGAGAGTCTGTTAAATAGTTGGCCACTTGGCAACCAAGAGATACCGCATTTGG 482  
 QY 412 CAATTTGAGGCGATGACTATTAGTCTCAAGATGAAAAATGGGCTCATCATCGAAA 471  
 DB 483 AGTTGCTAGCCAAAGGCTTGTAAATATGATGATGAAAAATGGGCAAGACACAGAGA 542

QY 472 ATCATTAAGCCCTAATTCTTTTCAATATGAGAAATCTTAAGTTGCTTGACAGTTGTTGAG 531  
 DB 543 ATCTTAATCTGCAATTTTCAACCATGAAAAATTAAGGGAGATGACCAATATTTTCAC 602  
 QY 532 AGTGTACTGATATGAGTGAATTAATGTCCTGATTAAGTTATCAAAAAACGGTGAAGTTAG 591  
 DB 603 TGCTGTATTAAGAAATGTTACTGATGAGATTAATTCATGCTTCTGAGGATCTCTGAG 662  
 QY 592 GTAGATGCTATGAGTGGTTTCAATTTTGACTGAAGATGTTATTAAGAAACAGCTTTT 651  
 DB 663 ATGATGTCGAGCTGAGTTCCAGATCTTCAAGATGTTATCTGAAAGATCGCTTT 722  
 QY 652 GGAATGATGATGAAGATGATGAGCAAGTTTGTGACTTCACTTCAACCAATAGCTCTT 711  
 DB 723 GGGAGCACTATCAAGAAAGGAGAGAGATTTTTCAGCTGCAAGAGAACTAGTGAACGC 782  
 QY 712 TGTGCTGAAGCTTTTCAAAAAAGCTTCAATTCCTGCTATGATTTTTCGACAAAGGG 771  
 DB 783 CTGATCAATCTGTTCAAGCAATATTTATCCAGGCTATGTTGTTCTTGCCACCAAAAC 842  
 QY 772 AATTGAAGTCTCGAAAGTTAGACAAGAGATTAAGAAAGTGTGTAAGCTGATAGAG 831  
 DB 843 AACGAAAGATGAGAGCAATGATGATGATGATGATGATGATGATGATGATGATGATG 902  
 QY 832 CGCGGAGACAAAACGCTATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 879  
 DB 903 AAG 962  
 QY 880 CGCGCGGAG 939  
 DB 963 CTGAGATCAAAACAGAT 1022  
 QY 940 ATTGTGAG 999  
 DB 1023 GTGATTAAG 1082  
 QY 1000 ACGTGAGACCATCTTGTGATTCATGACACCGAGAGAGAGAGAGAGAGAGAGAGAGAG 1059  
 DB 1083 ACTTGAGACCTATTTGTGCTAAGATGACATGACACCGAGAGAGAGAGAGAGAGAGAG 1142  
 QY 1060 GTCTTGAAGGTGTGGGCTCACGTGATCTCTTCAAGAGAGAGAGAGAGAGAGAGAG 1119  
 DB 1143 GTTTGAGGCACTTTGG---AAGAAACACACCGAGATTAAGTTAGCTTGAGCCGCTCAG 1199  
 QY 1120 ACGTTGAGATGATCTTGAAGAGAGCTTTTAAGTTGATTCACCAATGATGATGATGAT 1179  
 DB 1200 ACTATACCATGATTTCTACAGAGAGCTTGAAGTTGATGATGATGATGATGATGATGAT 1259  
 QY 1180 CGACGGCTTAATCGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1239  
 DB 1260 AAGAGAACTTAAG 1319  
 QY 1240 CTATCCCAATGATAGCGGCTCATCATGATGACCAAGCAATTTGGGTTATGACGTGAACGA 1299  
 DB 1320 CTCTTACCGGTATCTTATCATCATGATGATGATGATGATGATGATGATGATGATGATG 1379  
 QY 1300 TTTCAATCGAGCTGGTTTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1359  
 DB 1380 TTTCAATCGAG 1439  
 QY 1360 ATACGTTTGGCTGAGAGTTGATCATGATGATGATGATGATGATGATGATGATGATGAT 1419  
 DB 1440 TTTCCGTTGGAG 1499  
 QY 1420 AATTTGACACTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1479  
 DB 1500 AAGATGAGCTATGACCAATCTCTCAAGCTTCTGTTGAGAGCTCTGCAATCTCAAC 1559  
 QY 1480 CATGACCTTACCTGCTTATGTTGTTATCTTCAACAGTGTGACCAATCAACCTTC 1536  
 DB 1560 CACGCGCCCTTACACCGATTAACACTGACACCTTCAACAGTGTGCTCAGATTAAGGCTC 1616





Mon May 24 08:10:07 2004

us-09-992-901-1.rng

Page 8

```
DB      1324  GAAACGCGATTTCATGACCCACCATTCACCCAAAGCTCTCCCTCCGTTTCATCGGA 1393
QY      1378  GTTCGTACATGATGGTCAGAAATCTGTAATCTTCAGAGCCAAATGACATCGCTGTA 1437
DB      1384  CCAAGAGTTTCATGCGCAAAACTTGGCATGTCGAGCCAGACGCGTTCTCATATG 1443
QY      1438  ATGATCCACGCTTCACCTTCTCACTTGGCTCTTACTTATCAGATGACACTTACGCTCTT 1497
DB      1444  ATCCTTCAACAGTTCCAGCTTACGCTCTCCCGAGTATAGACACGCGCAGTTGATCAC 1503
QY      1498  ATGTGCTTATCTCTACATGTCACCATCA 1531
DB      1504  TTGATCTCTCTCCGCAATAGCGGCTACCGGTAA 1537

RESULT 7
AAC43009
ID      AAC43009 standard; DNA; 1712 BP.
XX
XX      AAC43009;
AC      17-OCT-2000 (first entry)
XX
XX      Arabidopsis thaliana DNA fragment SEQ ID NO: 37690.
DE
XX      Hybridisation assay; genetic mapping; gene expression control;
KW      protein identification; signal transduction pathway; metabolic pathway;
KW      promoter; termination sequence; ss.
XX
OS      Arabidopsis thaliana.
XX      EP1033405-A2.
XX
PD      06-SEP-2000.
XX
XX      25-FEB-2000; 2000EP-00301439.
PF
XX
PR      25-FEB-1999; 99US-0131825P.
PR      05-MAR-1999; 99US-013180P.
PR      09-MAR-1999; 99US-013548P.
PR      23-MAR-1999; 99US-0125788P.
PR      25-MAR-1999; 99US-0126264P.
PR      29-MAR-1999; 99US-0126785P.
PR      01-APR-1999; 99US-0127462P.
PR      06-APR-1999; 99US-0128234P.
PR      08-APR-1999; 99US-0128714P.
PR      16-APR-1999; 99US-0129845P.
PR      19-APR-1999; 99US-0130077P.
PR      21-APR-1999; 99US-0130449P.
PR      23-APR-1999; 99US-0130510P.
PR      28-APR-1999; 99US-0130891P.
PR      30-APR-1999; 99US-0131449P.
PR      30-APR-1999; 99US-0132048P.
PR      04-MAY-1999; 99US-0132407P.
PR      04-MAY-1999; 99US-0132484P.
PR      05-MAY-1999; 99US-0132485P.
PR      06-MAY-1999; 99US-0132486P.
PR      07-MAY-1999; 99US-0132487P.
PR      11-MAY-1999; 99US-0132863P.
PR      14-MAY-1999; 99US-0134218P.
PR      14-MAY-1999; 99US-0134219P.
PR      14-MAY-1999; 99US-0134221P.
PR      14-MAY-1999; 99US-0134379P.
PR      18-MAY-1999; 99US-0134768P.
PR      19-MAY-1999; 99US-0134941P.
PR      20-MAY-1999; 99US-0135124P.
PR      21-MAY-1999; 99US-0135353P.
PR      24-MAY-1999; 99US-0135629P.
PR      25-MAY-1999; 99US-0136021P.
PR      27-MAY-1999; 99US-0136392P.
PR      28-MAY-1999; 99US-0136782P.
PR      01-JUN-1999; 99US-0137222P.
PR      03-JUN-1999; 99US-0137528P.
PR      04-JUN-1999; 99US-0137502P.
PR      07-JUN-1999; 99US-0137724P.
PR      08-JUN-1999; 99US-0138094P.
PR      10-JUN-1999; 99US-0138540P.
PR      10-JUN-1999; 99US-0138847P.
PR      14-JUN-1999; 99US-0139119P.
PR      16-JUN-1999; 99US-0139432P.
PR      16-JUN-1999; 99US-0139453P.
PR      17-JUN-1999; 99US-0139492P.
PR      18-JUN-1999; 99US-0139454P.
PR      18-JUN-1999; 99US-0139455P.
PR      18-JUN-1999; 99US-0139456P.
PR      18-JUN-1999; 99US-0139457P.
PR      18-JUN-1999; 99US-0139458P.
PR      18-JUN-1999; 99US-0139459P.
PR      18-JUN-1999; 99US-0139460P.
PR      18-JUN-1999; 99US-0139461P.
PR      18-JUN-1999; 99US-0139462P.
PR      18-JUN-1999; 99US-0139463P.
PR      18-JUN-1999; 99US-0139750P.
PR      18-JUN-1999; 99US-0139763P.
PR      21-JUN-1999; 99US-0139817P.
PR      22-JUN-1999; 99US-0139899P.
PR      23-JUN-1999; 99US-0140353P.
PR      23-JUN-1999; 99US-0140354P.
PR      24-JUN-1999; 99US-0140695P.
PR      28-JUN-1999; 99US-0140823P.
PR      29-JUN-1999; 99US-0140991P.
PR      30-JUN-1999; 99US-0141287P.
PR      01-JUL-1999; 99US-0141842P.
PR      01-JUL-1999; 99US-0142154P.
PR      02-JUL-1999; 99US-0142055P.
PR      06-JUL-1999; 99US-0142390P.
PR      08-JUL-1999; 99US-0142803P.
PR      09-JUL-1999; 99US-0142920P.
PR      12-JUL-1999; 99US-0142977P.
PR      13-JUL-1999; 99US-0143542P.
PR      14-JUL-1999; 99US-0143624P.
PR      15-JUL-1999; 99US-0144005P.
PR      16-JUL-1999; 99US-0144086P.
PR      16-JUL-1999; 99US-0144086P.
PR      19-JUL-1999; 99US-0144325P.
PR      19-JUL-1999; 99US-0144331P.
PR      19-JUL-1999; 99US-0144332P.
PR      19-JUL-1999; 99US-0144333P.
PR      19-JUL-1999; 99US-0144334P.
PR      19-JUL-1999; 99US-0144335P.
PR      19-JUL-1999; 99US-0144336P.
PR      20-JUL-1999; 99US-0144632P.
PR      20-JUL-1999; 99US-0144684P.
PR      21-JUL-1999; 99US-0144814P.
PR      21-JUL-1999; 99US-0145086P.
PR      21-JUL-1999; 99US-0145088P.
PR      22-JUL-1999; 99US-0145085P.
PR      22-JUL-1999; 99US-0145087P.
PR      22-JUL-1999; 99US-0145089P.
PR      22-JUL-1999; 99US-0145192P.
PR      23-JUL-1999; 99US-0145145P.
PR      23-JUL-1999; 99US-0145218P.
PR      23-JUL-1999; 99US-0145224P.
PR      26-JUL-1999; 99US-0145276P.
PR      27-JUL-1999; 99US-0145913P.
PR      27-JUL-1999; 99US-0145918P.
PR      27-JUL-1999; 99US-0145919P.
PR      28-JUL-1999; 99US-0145951P.
PR      02-AUG-1999; 99US-0146386P.
PR      02-AUG-1999; 99US-0146389P.
PR      02-AUG-1999; 99US-0146389P.
PR      03-AUG-1999; 99US-0147038P.
PR      04-AUG-1999; 99US-0147204P.
PR      04-AUG-1999; 99US-0147302P.
PR      05-AUG-1999; 99US-0147192P.
```

PR 05-AUG-1999; 99US-0147260P.  
 PR 06-AUG-1999; 99US-0147303P.  
 PR 06-AUG-1999; 99US-0147416P.  
 PR 09-AUG-1999; 99US-0147493P.  
 PR 09-AUG-1999; 99US-0147935P.  
 PR 10-AUG-1999; 99US-0148171P.  
 PR 11-AUG-1999; 99US-0148319P.  
 PR 12-AUG-1999; 99US-0148565P.  
 PR 13-AUG-1999; 99US-0148684P.  
 PR 16-AUG-1999; 99US-0149368P.  
 PR 17-AUG-1999; 99US-0149175P.  
 PR 18-AUG-1999; 99US-0149426P.  
 PR 20-AUG-1999; 99US-0149722P.  
 PR 20-AUG-1999; 99US-0149723P.  
 PR 23-AUG-1999; 99US-0149929P.  
 PR 23-AUG-1999; 99US-0149902P.  
 PR 25-AUG-1999; 99US-0150566P.  
 PR 26-AUG-1999; 99US-0150884P.  
 PR 27-AUG-1999; 99US-0151065P.  
 PR 27-AUG-1999; 99US-0151066P.  
 PR 30-AUG-1999; 99US-0151080P.  
 PR 31-AUG-1999; 99US-0151438P.  
 PR 01-SEP-1999; 99US-0151930P.  
 PR 07-SEP-1999; 99US-0152363P.  
 PR 10-SEP-1999; 99US-0153070P.  
 PR 13-SEP-1999; 99US-0153758P.  
 PR 15-SEP-1999; 99US-0154038P.  
 PR 16-SEP-1999; 99US-0154039P.  
 PR 20-SEP-1999; 99US-0154739P.  
 PR 22-SEP-1999; 99US-0155139P.  
 PR 23-SEP-1999; 99US-0155486P.  
 PR 24-SEP-1999; 99US-0155558P.  
 PR 28-SEP-1999; 99US-0156458P.  
 PR 29-SEP-1999; 99US-0156596P.  
 PR 04-OCT-1999; 99US-0157117P.  
 PR 05-OCT-1999; 99US-0157753P.  
 PR 06-OCT-1999; 99US-0157865P.  
 PR 07-OCT-1999; 99US-0158029P.  
 PR 08-OCT-1999; 99US-0158232P.  
 PR 12-OCT-1999; 99US-0158369P.  
 PR 13-OCT-1999; 99US-0159293P.  
 PR 13-OCT-1999; 99US-0159284P.  
 PR 13-OCT-1999; 99US-0159295P.  
 PR 14-OCT-1999; 99US-0159329P.  
 PR 14-OCT-1999; 99US-0159330P.  
 PR 14-OCT-1999; 99US-0159331P.  
 PR 14-OCT-1999; 99US-0159637P.  
 PR 14-OCT-1999; 99US-0159638P.  
 PR 18-OCT-1999; 99US-0159584P.  
 PR 21-OCT-1999; 99US-0160741P.  
 PR 21-OCT-1999; 99US-0160767P.  
 PR 21-OCT-1999; 99US-0160768P.  
 PR 21-OCT-1999; 99US-0160770P.  
 PR 21-OCT-1999; 99US-0160814P.  
 PR 21-OCT-1999; 99US-0160815P.  
 PR 22-OCT-1999; 99US-0160980P.  
 PR 22-OCT-1999; 99US-0160981P.  
 PR 22-OCT-1999; 99US-0160989P.  
 PR 25-OCT-1999; 99US-0161404P.  
 PR 25-OCT-1999; 99US-0161405P.  
 PR 25-OCT-1999; 99US-0161359P.  
 PR 26-OCT-1999; 99US-0161360P.  
 PR 26-OCT-1999; 99US-0161361P.  
 PR 28-OCT-1999; 99US-0161950P.  
 PR 28-OCT-1999; 99US-0161952P.  
 PR 28-OCT-1999; 99US-0161953P.  
 PR 29-OCT-1999; 99US-0162142P.

Query Match

15.4%; Score 241.2; DB 3; Length 1712;

Best Local Similarity 49.2%; Pred. No. 1,8e-61;  
 Matches 754; Conservative 0; Mismatches 768; Indels 12; Gaps 4;

QY 10 GAAAGTACAGCTGGTTCATTCCAAAGGTTCTGTCTGTCTGTATCTTAGTCTGTA 69  
 DB 4 GAAGCTTAAAGCACAACTAACTCTTAACAGTGTCTCTCTCTCTCTCTCTCTCAAG 63

QY 70 ATAGTGAAGGTATGCTCTGTATAGTGTAGACCAAGAAAGATTGAAGACATTTCT 129  
 DB 64 ATATGAAGAGCTTGTGATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCAAG 123

QY 130 AAACAGGAATTCAGAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 189  
 DB 124 AAACAGGAATTCAGAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 183

QY 150 GAAATGATGCTTAAAGCT--TCTTCTCATCTTATGCTCTCTCTCTCTCTCTCTCT 246  
 DB 184 AAGATGAAGAAAGAGCTGACCTTGTGTCTGATCCCAATTCATGATATTTCCCT 243

QY 247 AGAGTCT 306  
 DB 244 CGTGTCT 303

QY 307 TTGGTCCAACTTCCGCTTACGGTACCGATCTCTGATTTGATCAGAGATCTTCTC- 365  
 DB 304 ACTGGAACAAACCACTATATATCATCTCTCAATCATGAGCTAGCGAAACAGGCTTGTG 363

QY 366 --TAGCTGAGTCTACGAGAAAGATGAAGCTCACCCCTTGTGTAACACTGAGGC 423  
 DB 364 AGCAAGTTCGTTCTACTATTATACCGGTGAACAGCTGAGGCTTCTCATCTTCTGCT 423

QY 424 GATGACTACTTACTCTCAAGGTGAAGAAATGGGCTCATCTCAAGAAATCTATGACCT 483  
 DB 424 AAAGACTTTCCTTTATCCAAAGGTGATGATGATGCGCATAGACGAATCTTGAACCT 483

QY 484 ACTTTGATATGAGATCTTATGTTGCTTGTACAGGTTGTGAAGTGTGATCTAT 543  
 DB 484 GCTTTCCTCAGATCGGCTGACCGATGACCAACCGATGAGATGACCTTACCTTACG 543

QY 544 ATGTGTGATTAATGTCGATTAATGATCA--GAAACGGTGAAGTTGAGTATGATC 600  
 DB 544 ATATTGAGAGTGTGAAGAAACAGAGAAATGTGAAGTGTATCAAGTATGATTA 603

QY 601 TATGATGCTTTCGATTTGATGAAAGTATTAAGTAAAGCTTGTGAAGTATGAC 660  
 DB 604 AGCAAGATTCCTAATGACCTGACCTGACATATGCACTACCTGCTTGTGAAGTAT 663

QY 661 TATGAAGTGTTCAGACGTTTTCGACTTCAAGCTCAACAAATGCTTGTGTGCTGA 720  
 DB 664 TATGCGAAGGATCGAATGTGTAGTGCACAAACAGACTTGAGAGATATATATG 723

QY 721 GCTTTTCAAAAGCTTCATCTCTGCTATTAATTTTTCGACAAGGGAATTTGAG 780  
 DB 724 TCTCTCACTAACGTTTATCTCCCGAAGTCAATTAATCTTCTTACGCTTCAACTTAA 783

QY 781 TCTCGAAGTTAGCAAGAGATTAAGAACTGTTGTGAAGCTATAGA--GCGGCGG 837  
 DB 784 CTATGGAGGCTCCATTAAGAAAGTGAAGATCATCAAAAGATCATATTAATTAAGGCTA 843

QY 848 AGACAAACGCTATAGTGAAGAGGGAAGATGAAGACCGGCGCGCAAGATTTG 897  
 DB 844 AATATCAAAATTAAGACTTATGCTATGGGAGAGATCTTCTAGGGCTCATGTGACTCT 903

QY 898 TTGGATTAATGATTCAGGCAAGAAATGTGACGTTCAAGACATTTGTGAAGATGTA 957  
 DB 904 GCGAAATCTAACGATAGAGAAAGATGAGATGAGATCTATAGAGATGACAG 963

QY 958 AGCTTTTCTTCTGCGGGAAGAAACAGCAACTTAACTGTGACGTCGACCACTTGT 1017  
 DB 964 AATTTCTATAGAGGCAAGATCTACTTCAATTTATGACATGACATGATGTA 1023

QY 1018 CTATCCATGACCCGAGGTGCGAGGCAAGACATGATGAGTCTCTAGGCTTGTGCGC 1077

Db 1024 CTAACTTGCACCAAGGTGGCGAAGAACTCAGAGAAAGGTTTAAATGATGTGT 1083  
 QY 1078 TCACGTGATGTCCTTCCACAGAGCCATGTCTTAAGCTTAAAGCTTGAATGATCTTG 1137  
 Db 1084 AAAGATTAATCCCGATACAGACACCTTCTTAATCTCAATTAATGATGAATGTGTG 1143  
 QY 1138 AACGATCTTAAAGGTGTATCCCAATAGTACCTCATGTGAGCGCTAAATGCGAT 1197  
 Db 1144 ATGGATCGCTTCTCTTACGAGCCGTGATTAATTTCCGAGAGGCAACAAGAT 1203  
 QY 1198 GTGAAGCTAGAGAGGTACCAAAATCCATGTGGCAAGGAGCTTCAATCCCAATGAG 1257  
 Db 1204 ATGAAGGTAGGACCTTGGAGATCCCAAGGCGCAAGCATATATACCTGTTACTGAAG 1263  
 QY 1258 GTCCATCATGACCAAGCATTTGGGGTAAATGAGGTGAAGCAATTCATCACTGAGCTT 1317  
 Db 1264 ATGCACAGAGACAGAGCCATATGGGGCGAAGACCGCAATTCATCTCTTGAGATTC 1323  
 QY 1318 GCGATGAGATGCGCGGTGTCGCAAAACCCCGTTGGCTTCAATACCGTTGGCGCTCGGA 1377  
 Db 1324 GAAACCGCATTTCTCAAGCCACCATTCACCAACGCTCTCTCCCTCCCTCAATCGGA 1383  
 QY 1378 GTTCGATCATGATGTGTGATGATCTTGTATCTTCAAGCCCAATTCACCTGCTGTA 1437  
 Db 1384 CCAGAGCTTGCATTTGCCAAATCTTGCATGTGTGAGAGCCCAAGCCGTTCTCACTATG 1443  
 QY 1438 ATGATCAACGCTTCACTTCACTTGTGCTCTTCACTTATCAGATGCACTTACCGTCTT 1497  
 Db 1444 ATCTTCAACAGTTTCAGAGCTTACGCTCTCCCGAGATPAGACACGCGATGTATGAC 1503  
 QY 1498 ATGTCTTATCTCTCAACATGTGCAACCAATCA 1531  
 Db 1504 TTGATCTCTTCCGCAATAGGCTTACCGGTA 1537

RESULT 8  
 ADA8311  
 ID ADA8311 standard; DNA; 1365 BP.  
 XX  
 AC ADA8311;  
 XX  
 DT 20-NOV-2003 (first entry)  
 XX  
 DE Rice gene conferring disease resistance in plants.  
 XX  
 KW disease resistance; pathogen tolerance; plant pathogen; de; gene; plant.  
 XX  
 OS *Oryza sativa*.  
 XX  
 PN WO200300906-A2.  
 XX  
 PD 03-JAN-2003.  
 XX  
 PF 21-JUN-2002; 2002WO-IB002453.  
 XX  
 PR 22-JUN-2001; 2001US-0300112P.  
 PR 26-SEP-2001; 2001US-0352277P.  
 PR 22-MAR-2002; 2002US-0365535P.  
 XX  
 PA (SYGN) SYNGENTA PARTICIPATIONS AG.  
 PI Glazebrook J, Briggs S, Cooper B, Goff SA, Moughamer T,  
 PI Katagiri F, Kreps J, Provart N, Rieke D, Zhu T;  
 PI  
 DR WPI: 2003-184052/18.  
 DR P-PSDB: ADA8312.  
 XX  
 PT New polynucleotide comprising a plant nucleotide sequence having an open  
 PT reading frame that encodes a polypeptide associated with disease  
 PT resistance, useful for conferring resistance or tolerance to a plant  
 PT pathogen.  
 XX  
 PS Claim 1; SEQ ID NO 381; 299bp; English.

XX The invention relates to a novel isolated polynucleotide comprising a  
 CC plant nucleotide sequence having an open reading frame that encodes a  
 CC polypeptide associated with disease resistance or its fragment having  
 CC substantially the same activity as the full-length polypeptide. The  
 CC polynucleotide of the invention is useful for conferring resistance or  
 CC tolerance to a plant pathogen. The present sequence represents a gene  
 CC conferring disease resistance used in the invention.  
 XX  
 SQ Sequence 1365 BP; 315 A; 341 C; 409 G; 299 T; 0 U; 1 Other;  
 Query Match 13.2%; Score 206.2; DB 8; Length 1365;  
 Best Local Similarity 50.4%; Pred.No.5.5e-51;  
 Matches 639; Conservative 0; Mismatches 613; Indels 15; Gaps 5;  
 QY 292 AATTTCTGTTGGTTGGTTCACACTTCCGTTACAGGTACCGCATCTGTATTGATC 351  
 Db 94 AATTTCTGTTGGTTGGTTCACACTTCCGTTACAGGTACCGCATCTGTATTGATC 351  
 QY 352 AGAGAGATCTTCTTAA--GTCTGATTTACAGAGAGATGAGGTCACTCCCTTGTGTT 408  
 Db 154 TGGCAGATTTGTCAAGCCGAGCGGATATACCCCAAGACCTGACGAACCTTCACTTT 213  
 QY 409 AAACAATTGAAGCGAGATGACACTTACTAGTCTCAAAAGGTGAAAATGGGCTCATCTGA 468  
 Db 214 GTGCGTTACTTGGCAAGGGGCTTGTCTCACTGACGCGAGATGAGAGAGCGCACCGC 273  
 QY 469 AAATCATTAAGCCCTTACTTTCATATGAGAGATCTTAAGTGTCTTGAACAGTTGTGTTG 528  
 Db 274 AAGGTGTCCACCGCGCTTATACATGACAGCTCAAGATGATGACATGACATGATCT 333  
 QY 529 AAGGTGTACATGATGTGTGATTAATGTCTCCATTAAGTATCAAAAAGGTG--AA 585  
 Db 334 GACTCTCTCGGTCATGATGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 393  
 QY 586 GTTGAAGTATGATGTGTGATGATGTGTGATGATGATGATGATGATGATGATGATGATG 645  
 Db 394 GTTGAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 453  
 QY 646 GCTTTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 705  
 Db 454 GCTTTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 513  
 QY 706 CTTCCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 765  
 Db 514 TTCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 573  
 QY 766 AAGAGATTTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 825  
 Db 574 ATGAATAAATTCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 633  
 QY 826 ATAGAG--CGGCGAGCAAAACGCTATGATGATGATGATGATGATGATGATGATGATG 882  
 Db 634 ATCAAGACCCGCAATGCAACCAAGATGATGATGATGATGATGATGATGATGATGATG 693  
 QY 883 GCGGAGAGATTT--GTGGGATTAATGATGATGATGATGATGATGATGATGATGATG 939  
 Db 694 ATGCTGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 753  
 QY 940 ATTGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 999  
 Db 754 ATCATTTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 813  
 QY 1000 AAGTGAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1059  
 Db 814 AACTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 873  
 QY 1060 GTCTCAGAGTGTGCGGCTACGATGATGATGATGATGATGATGATGATGATGATGATG 1119  
 Db 874 ATTTGCAATGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 930  
 QY 1120 ACGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1179

Db 931 ATGCAACATGTTCTCTCTAGAGACCCCTAAGGCTGAGACCCCTCTCTCACTCATACG 990  
 Qy 1180 CGAGCGCGTAAATGCGATGTAAGCTAGAGGCTAAGATCCCATGTGCGACGAGCTT 1239  
 Db 991 AGAAGGTTGACATGATTAATGAGCTGCGGCGATCAAGATGCTGAGGGTGCCTATTG 1050  
 Qy 1240 CTAATCCCAATCATAGGCGTCCATCATGACCAAGCCATTGGGGTAAATGACGTGAAGAA 1239  
 Db 1051 ACGATCCCGATGCGACAAATTCACCTGACAAAGAAAGTGTGGGGGAGAGATCGGATGAG 1110  
 Qy 1300 TTCAATCCAGCTCGTTTGGCGAGTNGAGTCCCGCTGCTGCCAAACACCCCTGTGCTTC 1359  
 Db 1111 TTTAGGCGCGAGAGTTGAGAAATGGGGTACAGAGGACCGAAGACCCCAAGCGCACTA 1170  
 Qy 1360 ATACGTTTGGCTCGAGGTTCTGATCATGCTTGTGCAAGATCTGTACTTACTTACGAGCC 1419  
 Db 1171 CTCTCTCTCTCCAGCGGACCGAGGTATGATGATGGGACAGACTTGGCAATGATGAGGCC 1230  
 Qy 1420 AAATTGACATCGCTGTATGATCCAAAGCTTCACTTTCATCTGCTCTTACTATGAG 1479  
 Db 1231 AAGGCTCATTTGCCATGATCTCAAGAGGTTCTCTTCACTTGTCCCAAGTACGTC 1290  
 Qy 1480 CATGACCTTACCGTCTTATGCTTATCTTATCTCAACATGTCACCAATCAGCTTCCGG 1539  
 Db 1291 CATGACCCCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1350  
 Qy 1540 AGATTGA 1546  
 Db 1351 AGCTTCA 1357  
 RESULT 9  
 ID AB212574 standard; DNA, 1539 BP.  
 AC AB212574;  
 DT 21-JAN-2003 (first entry)  
 XX Arabidopsis thaliana stress regulated gene SHQ ID NO 379.  
 DE Arabidopsis thaliana; plant; gene; stress; transgenic; ds.  
 KM Arabidopsis thaliana; plant; gene; stress; transgenic; ds.  
 OS Arabidopsis thaliana.  
 PN WO200216655-A2.  
 XX 28-FEB-2002.  
 PD 24-AUG-2001; 2001WO-US026685.  
 PF 24-AUG-2001; 2000US-0227866P.  
 XX 26-JAN-2001; 2001US-0264647P.  
 PR 22-JUN-2001; 2001US-0300111P.  
 XX (SCRI ) SCRIPPS RES INST.  
 PA (SYGN ) SYNGENTA PARTICIPATIONS AG.  
 XX Harper JF, Kreps J, Wang X, Zhu T;  
 PI WPI; 2002-304127/34.  
 DR WPI; 2002-304127/34.  
 XX Identifying a stress condition to which a plant cell has been exposed and  
 PT producing plants with increased tolerance to these abiotic stresses.  
 XX Claim 144; SEQ ID NO 379; 577bp + Sequence Listing; English.  
 BS The invention relates to identifying a stress condition to which a plant  
 CC cell has been exposed, comprising: (a) contacting nucleic acid  
 CC representative of expressed polynucleotides in the plant cell with an  
 CC array of probes representative of the plant cell genome; and (b)  
 CC detecting a profile of expressed polynucleotides in the plant cell  
 CC characteristic of a stress response. The method is useful in the

CC production of transgenic plants, cells and seeds and in producing plants  
 CC with increased tolerance to abiotic stress. The present sequence is that  
 CC of an Arabidopsis thaliana stress regulated gene (AB212574) used  
 CC in methods of the invention. Note: The sequence data for this patent is  
 CC not represented in the printed specification but is based on sequence  
 CC information supplied to Derwent by the European Patent Office  
 XX  
 SQ Sequence 1539 BP; 439 A; 344 C; 349 G; 407 T; 0 U; 0 Other;  
 Query Match 12.5%; Score 194.8; DB 6; Length 1539;  
 Best Local Similarity 48.7%; Pred. No. 1.6e-47;  
 Matches 702; Conservative 0; Mismatches 707; Indels 33; Gaps 5;  
 Qy 100 AGACCAAGAAAGATTGAAGAAATTTCTTAAACAAGAAATTCAGAGCTCTCTTATCAT 159  
 Db 94 ACACCAAGCAATCAAGAAATTTATGAAAGCTCAAGCAATCAAGGCTTAAACCCGT 153  
 Qy 160 TTCTTCATCGAAATGTTAAAGACTGTGAAATGATGCTTAAAGCTTCTCTCATCT 219  
 Db 154 CTCTCACCAGAAATCATGATGATGATGATGATGATGATGATGATGATGATGATGATG 213  
 Qy 220 ATGCTT---TCTCAAAATATCTTCTAGAGTTCTCTCTTTTACCATCATGAGA 276  
 Db 214 TGCTCTTCATTCACCAACATCATGCTCCAGGCTTCTCTTATGATGATGATGATGATG 273  
 Qy 277 AAAATCTACGGTCTACCTTTCTGTTTGGTTGGTCCAACTTCCGTTAAGGTAAGCC 336  
 Db 274 AAACATATGGAAAGATTTATATATGAGAGGAGACGAGAACGAGACTATGCTTAAAG 333  
 Qy 337 GATCTGATTTGATCAGAGATCTTCTTAAAGTCTGATGATGATGATGATGATGATGATG 396  
 Db 334 GAGACGAGATGATTTAAAGATGCTTAAAGAGACATCCGCTCACCGGAAATCATG 393  
 Qy 397 -----CACCTTGGTTTAAACAATGAAAGGAGATGACTTATGCTCAAGGTA 450  
 Db 394 CTGCAACAACAAGAACTTAAAGGTTTATATGAGGCTGAGGCTTCTCAAGGCAAGG 453  
 Qy 451 AAATGGCTCATCATGCAAAATCATTAAGCTTCTTCTTCAATGAGATCTTAAAGTTG 510  
 Db 454 GCTTGCAACACCAACGTCATGACATGACGCTCCGCTTTTCAAGTATAGGCTCAAAG 513  
 Qy 511 CTGTACCAAGTTGTTGAAAGTGTGATGATGATGATGATGATGATGATGATGATGATG 570  
 Db 514 TACGCTAAGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 561  
 Qy 571 TCAGAAACGCTGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 630  
 Db 562 AGAAAGAGGTTGAGAAAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 621  
 Qy 631 GTTATTAATGAACAGCTTTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 690  
 Db 622 AATAATGAGAGAGGAAATTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 681  
 Qy 691 CAAGCTCAACAATGCTTCTTGTGCTGAAGCTTTTCAAAAGTCTTCACTCTGCTAT 750  
 Db 682 CTCACTGATTTCAACAGCTTTTGTGCTCAAGCACTGCACTCTGTTTCCCGTAGC 741  
 Qy 751 AGATTTTTCGAAAGAGGAAATTTGAATCTCGGAATGTAGACAAAGAAATGAAGAG 810  
 Db 742 CGGTTTCACTTACCAATACCAATGAGAAATTAAGTCTGTGAAACGGAAGGAGAGCT 801  
 Qy 811 TCGTTGTTGAAGCTGATGAGCGCGGAGACAA---ACGCTATGATGAGAGAGGAA 867  
 Db 802 CTTTGAATGAGATTAATGAGACGCAAGAAAGACAGTGTGAGATCGGTGAGTAGCTCA 861  
 Qy 868 GAATGTAAGAGCCGCGGAGAGATTTGTTGGATTAATGATTAAGGCAAAAGATG 927  
 Db 862 TACGGGATGACCTTTAGGCTCTTTTGAACAAGATGATGACAAAGAACCTT 921  
 Qy 928 ACGTTCAAGCATTTGAGAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 987  
 Db 922 AATGTTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 981

```
QY 988 TCTAATCTGCTGACGTGACGACCATCTTGTCTATCATGACCCCGAGTGGCAAGCCAAA 1047
DB 982 TCTTCTCTCTCACTGACATCTCATGCTCTTGTCTATCATGACCCATGCGAGGACAT 1041
QY 1048 GCAGGTGATGAGTCTCTGAGGTGCGGGCTGCGAGTATGCTCTGACCAAGGACATGTC 1107
DB 1042 GTCCCGATGAGTCTGCGCAATTTGCGCCAGATGATGTCCTTCCCTGTAAGAGCTC 1101
QY 1108 GTTAAGCTTAAACGTTGATGATCTTAAACGATCTTAAAGTTGATTCACCAATA 1167
DB 1102 TCAAGCTTACTCTTAAACAAAGTATTAACAGATCATTAAGACTTTACCTCTGTC 1161
QY 1166 GTTAGCTAGATTCACGCGGTAAATCGATGTGAGGTAGAGGTACAAATCCAGT 1227
DB 1162 AACTTTTACCAAGATGCACTTGAAGACATTAACCTAGGTACTATCTCTTAA 1221
QY 1228 GGCACGAGCTTCTAATCCATCATAGCGTCCATCATGACCAAGCAATTTGGGTAT 1287
DB 1222 GCTTATCTATTGATCTCTGCTCTGCGATCATCATATGATTAATGATGAGGCGGA 1281
QY 1288 GACGTGAAGATTAATCCAGCTCGGTTTGGGATGAGTGGCGGTGCGCAAC 1347
DB 1282 GATGCGAAGATTAATCCAGCTCTGAGATTTACTAGAGTTTGCATCTAGTGCAC 1340
QY 1348 CCCGTTGGCTTCAACGTTGGCTGAGTGTGACATGATGATGATGATGATGATGAT 1407
DB 1341 TTTATCTCTTTTGGCGTGGACCTGAGATGATGATGATGATGATGATGATGATGAT 1392
QY 1408 ATACTTCAAGCAATTAATGACCTGCTGATGATGATGATGATGATGATGATGATGAT 1467
DB 1393 ATGATGAGGCGAATTAATGATGATGATGATGATGATGATGATGATGATGATGATG 1452
QY 1468 CCTACTTACGATGACCTGACCTGCTTATGTTGCTTATGCTTATGCTTATGCTTATG 1527
DB 1453 GAGAACTATGACATGCTCTATGTTGCTTATGCTTATGCTTATGCTTATGCTTATG 1512
QY 1528 AT 1529
DB 1513 TT 1514

RESULT 10
AAC41990
ID AAC41990 standard; DNA; 1695 BP.
XX AAC41990;
AC AAC41990;
DT 17-OCT-2000 (first entry)
XX
XX Arabidopsis thaliana DNA fragment SEQ ID NO: 33879.
XX
XX Hybridization assay; genetic mapping; gene expression control;
XX protein identification; signal transduction pathway; metabolic pathway;
XX promoter; termination sequence; ss.
OS Arabidopsis thaliana.
XX
XX EP1033405-A2.
XX
XX 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-00301439.
XX
XX 25-FEB-1999; 99US-0121825P.
XX 05-MAR-1999; 99US-0123180P.
XX 09-MAR-1999; 99US-0123548P.
XX 23-MAR-1999; 99US-0125788P.
XX 25-MAR-1999; 99US-0126264P.
XX 29-MAR-1999; 99US-0126785P.
XX 01-APR-1999; 99US-0127462P.
XX 06-APR-1999; 99US-0128234P.
XX 08-APR-1999; 99US-0128714P.
XX 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 25-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 04-MAY-1999; 99US-0132407P.
PR 05-MAY-1999; 99US-0132484P.
PR 06-MAY-1999; 99US-0132486P.
PR 07-MAY-1999; 99US-0132487P.
PR 11-MAY-1999; 99US-0132566P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134368P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136329P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138647P.
PR 11-JUN-1999; 99US-0138119P.
PR 16-JUN-1999; 99US-0138452P.
PR 16-JUN-1999; 99US-0138453P.
PR 17-JUN-1999; 99US-0138492P.
PR 18-JUN-1999; 99US-0138454P.
PR 18-JUN-1999; 99US-0138455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139893P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 13-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 16-JUL-1999; 99US-0144325P.
PR 18-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
```

```

PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144332P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144684P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-014512P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147322P.
PR 05-AUG-1999; 99US-0147280P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156488P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157733P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159283P.
PR 13-OCT-1999; 99US-0159284P.
PR 14-OCT-1999; 99US-0159285P.
PR 14-OCT-1999; 99US-0159329P.

PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160780P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160982P.
PR 25-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161922P.
PR 28-OCT-1999; 99US-0161933P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 11.9%; Score 185.4; DB 3; Length 1695;
Best Local Similarity 49.0%; Pred. No. 1.1e-44;
Matches 751; Conservative 0; Mismatches 691; Indels 91; Gaps 6;

QY 90 GTTATGCTGAGACAGAAAGATTGAAGACATTTCTTAACAAGAAATTCGAGTTC 149
DB 139 GCTTTGCTTAACCAAGATGCTTGAAGTTCCTGAAGACAGAGCTTCCGGGAAAC 198
QY 150 TCTTATCATTTCTTCAATCGAAATGTTAAGACTTTGTTGAATCATCTTAAGCTTC 209
DB 159 TCTTATCAAGCCCTTCTCGTATTTGAAGAAATTTTACATGCTGAGAGCAAG 258
QY 210 TTTCAATCCATGCTTCTCTCAAAATTTCTTCAAGATTCCTTTTACATCA 269
DB 259 ATCCAAACCATCAACCTTACAGATATACACCAAGATTCGCTTATCCGTTGA 318
QY 270 CTGAGAAAATCTACGCTCTACATTTCTGTTGTTGGTCCACTTCCGGTTAC 329
DB 319 AATGCTCAAGCTCATGAGAGACTTCTTACATGTTGAGACCATCAACAATCAC 378
QY 330 GGTAGCCGATCCGATTTGATCAGAGATCTTCTTAAGTCTGAATTTACAGAGAA 389
DB 379 CATATGATCTTGAAGAAATCAAGAGATTTCAAGAAATTTGATTTCAAGAGGC 438
QY 390 TGAAGCTACCTTTGTTAACAATTGAAGGCGATGACTACTTACTTAAGTCAAGTCA 449
DB 439 ACATAGCTTCTCTGAGAGATTTAAT--AGCGCTGAGCTGTTAGTATGATGAGTGA-495
QY 450 AAAATGGCTCATCATGAAGAAATCATTAAGCCCTTACTTTTATGAGAAATCTTAAGT 509
DB 456 TAAATGACAAAGACCGAAGATCATCAACCGGCTTTCCACCTTGAAGAAATCAAAA 555
QY 510 GCTTGATCAAGTTGTGTAAGATGATGATGATGATGATGATGATGATGATGATGAT 569
DB 556 TATGATACCTGCTTCCACCAAGAGCTGACGAGATGTTGCGAATGGAACAATTAAGT 615
QY 570 ATCAGAAAAGGTGA--CTTAGGTATAGTCTATGATGATGATGATGATGATGATGAT 626
DB 616 GACGATTAACAGATCACTGATGATGATGATGATGATGATGATGATGATGATGATGAT 675
QY 627 AGATGTTATTAAGTGAACAGCTTTTGAAGTACTATGAAGATGTCGACAGTTTTCG 686
DB 676 AGATGATATCTCCGATCTCATTTGACACACTTCAAGAGGCGACAGATTAATTTGA 735
QY 687 ACTTCAAGCTCAAGAAATGCTTTTGTGCTGAAGCTTTTCAAAAGTCTTATTCCTGG 746
DB 736 GCTCAAGCGGAATTAACAGACTCATCAAGCTTATTAAGCTTTTGGAAAGCTATATCCCTGG 795

```

QY 747 CTATAGATTTTCCGACGAGGAAATTGAAGTCTCGAAGTTAGACAGAGATPAG 806  
 DB 796 AATGCTATCTCTCTCAAAAGGCTTATAGAGG---ATGAAAGCGAGCGTAAAGAAAT 852  
 QY 807 GAGTCGTTGTTGAAGCTGATAGACCGCGGAGCAAAAGCTATATGAGAAAGGGA 866  
 DB 853 CAATTTTACTGAGAGGATCTTTAAACAAAGGCTTAAAGGAGGACCTGGGAGACA 912  
 QY 867 AGAATGTAGAGAGCGCGGCGGAGATTTGTTGGATTATGATTCAGCAAAAGAT- 924  
 DB 913 CCAAGTGAAGATTTGCTGGGATATCTTCTGATTCGATTTGGGCAAAACAAAGAAAT 972  
 QY 925 ---GTGACGCTTACGAGCATTTGAGAGAGTGTAAAGCTTTTCTGCGCGGAAACAG 981  
 DB 973 GGAATGATGACCGAGGAACTGATGAGAGTGTAAAGTGTCTTATTTCTGCTGGGAGAG 1032  
 QY 982 ACAATCTTAACTCTGACGAGCGAGACCATCTTGTCTATCCATGACCGGAGTGGAG 1041  
 DB 1033 ACTACTAGAGTACTCTGCTGTTGAGCAATGGTGTGTTAAGCCACACCAAGATTGGAG 1092  
 QY 1042 GCCAAAGCACTGATGAGTCTCTAGGGTCTGGGC-----TCAGTGTATGCTCTAC 1095  
 DB 1093 GCTGCTGCAAGAGAGAGTGAACCAAGTTTGGGAGTAAAGAACCTGATCAGAGAGC 1152  
 QY 1096 AAGGACCATGTC----- 1107  
 DB 1153 TTGAACCAAGCTCAAAAGTTGTATGATTAACCGAACCAAAACCTTCAAGTGAAGTCT 1212  
 QY 1108 -----GTAACTTAAACGTTGAGTATGATCTTGAACGAGT 1144  
 DB 1213 TTGCTGATTTGAAGCTGAAGATTTATCTCTCAACAGATGACGATGATTTATGAG 1272  
 QY 1145 CTTTAAAGTTGATCTCAACCAATAGTATGATCTGACGCGCTAATGAGTGAAGC 1204  
 DB 1273 TCTTAAAGCTATATCTCTCAAGTATGATCTGACGCGCTAATGAGTGAAGC 1332  
 QY 1205 TAGAGGCTAATAATCCCATGATGACGAGCTTCTTATCCCATGATGAGGCTCATC 1284  
 DB 1333 TAGTGAACCTGACACTACAGGCGCGCTTCAAGTGAAGTATGATGATCTGATTCAC 1392  
 QY 1265 ATGACCAAGCATTTGGGCTATGACGCTGAACCAATTCATCCAGCTGCTTGGAGATG 1324  
 DB 1393 GCGACAGGAGCTGCTGGGAAAGCATGACAGGAGTTCAAGCCAGATGATTCAGAGAG 1452  
 QY 1325 GAGTCCGCGCTGCTGCCAAGACCCCGTGGCTTCAATCCGTTGGCTCTCGAGATGCTA 1384  
 DB 1453 GTCTCTCAAAAGCAAAAGAACCAAGCTCTCTTCTTCTTGGGAGGAGCCGAGAGA 1512  
 QY 1385 CATGATGCTGACATCTTGTCTATCTGACGCAAAATTCACATCGCTGATATGATCC 1444  
 DB 1513 TCTGATGCGCAAAATTTTGGACATGCTTGAAGGAAAGATGCAATGACATGATTCAC 1572  
 QY 1445 AAGCTTCACTTCACTGAGCTCTTCACTTATGACATGACATCCCTTATGATGCTC 1504  
 DB 1573 GGAATTTCTCTTGAAGCTTCTCTTATGATGACAGGCTTACACAGCTCTCACTA 1632  
 QY 1505 TTTATCTCAACATGTCACCAATCACTTCC 1537  
 DB 1633 CTCACCAAGTTGGTGTCTCTCTCACTTTCG 1665

## RESULT 11

AAA9071  
 ID AAA9071 standard; DNA; 588 BP.

AAA9071;

18-JAN-2001 (first entry)

Arabidopsis thaliana basal promoter sequence SEQ ID NO:16.  
 Arabidopsis thaliana; basal promoter; cytochrome P450; CYP72B1; plant;  
 brassinosteroid signaling; brassinosteroid synthesis; brassinolide; ds.

XX Arabidopsis thaliana.  
 OS Arabidopsis thaliana.  
 PN WO200055302-A2.  
 PD 21-SEP-2000.  
 PF 16-MAR-2000; 2000WO-US006915.  
 XX 16-MAR-1999; 99US-0124570P.  
 PR 14-DEC-1999; 99US-0170931P.  
 PR 20-DEC-1999; 99US-0172832P.  
 XX (SALK ) SALK INST BIOLOGICAL STUDIES.  
 PA Neff MW, Chory J;  
 PI WPI; 2000-638195/61.  
 DR Transgenic plants having modulated brassinolide synthesis resulting in  
 PT insect resistance, dwarfism and darker-green foliage compared with wild-  
 PT type plants, have nucleic acid encoding BAS1 polypeptide in its genome.  
 XX Claim 79; Fig 1C; 104p; English.  
 PS The present invention describes a genetically modified plant (1)  
 CC comprising at least one exogenous nucleic acid sequence encoding a BAS1  
 CC polypeptide, homologue or functional fragment, in its genome or at least  
 CC one regulatory sequence that modulated expression of endogenous basal gene,  
 CC homologue or functional fragment, and which is characterized as having  
 CC modulated brassinolide activity or synthesis. The basal gene encodes a  
 CC cytochrome P450 (CYP72B1), which has a role in brassinosteroid signaling  
 CC or synthesis. Overexpression of the basal gene in plants causes a dark  
 CC green, dwarf phenotype which mimics plants that have low levels of the  
 CC plant hormone, brassinolide. Overexpression of the basal gene also  
 CC increases resistance to insects in plants. The present sequence  
 CC represents the Arabidopsis basal promoter sequence, which is used in the  
 CC exemplification of the present invention  
 XX  
 SQ Sequence 588 BP; 193 A; 121 C; 99 G; 173 T; 0 U; 2 Other;  
 Query Match 11.8%; Score 184; DB 3; Length 588;  
 Best local similarity 99.0%; Pred. No. 1.7e-44;  
 Matches 206; Conservative 0; Mismatches 0; Indels 2; Gaps 2;  
 QY 1 ATGAGGAAAGAAAGTGAAGCTGTTATTCCTCAAGGTTCTTCTGCTGTAATCTTA 60  
 DB 362 ATGAGGAAAGAAAGTGAAGCTGTTATTCCTCAAGGTTCTTCTGCTGTAATCTTA 441  
 QY 61 AGT-CTGTAAATGAGAAAGGTATGCTCTGTTATGATGAGAGCAAGAAAGATTGAGA 119  
 DB 442 AGTCTTGTAAATGAGAAAGGTATGCTCTGTTATGATGAGAGCAAGAAAGATTGAGA 501  
 QY 120 AATTTCTTAAACAGAAATTCAGAGTCTCTTATCTTCTTATGAGAAAGTTAA 179  
 DB 502 AATTTCTTAAACAGAAATTCAGAGTCTCTTATCTTCTTATGAGAAAGTTAA 561  
 QY 180 AGAATTTGTAATGATGATCTTAAAGCT 207  
 DB 562 AGAATTTGTT-GAATGATGCTTAAAGCT 588

## RESULT 12

AEN98320/c  
 ID AEN98320 standard; DNA; 1124 BP.

AEN98320;

01-AUG-2002 (first entry)

Arabidopsis thaliana expressed polynucleotide SEQ ID NO 88.  
 Arabidopsis thaliana; plant; insecticide; fungicide; transgenic; stress;

XX disease crop; thale cress; tolerance factor; insect; pathogen;  
KW nutrition; ds.  
XX Arabidopsis thaliana.  
XX US2002023281-A1.  
XX 21-FEB-2002.  
XX 26-JAN-2001; 2001US-00770445.  
XX 27-JAN-2000; 2000US-0178472P.  
XX (GORL/) GORLACH J.  
PA (ANY/) AN Y.  
PA (HAMI/) HAMILTON C M.  
PA (PRIC/) PRICE J L.  
PA (RAIN/) RAINES T M.  
PA (YUY/) YU Y.  
PA (BAME/) BAMEKA J G.  
PA (PAGE/) PAGE A.  
PA (MATH/) MATHEW A V.  
PA (LEDF/) LEDFORD B L.  
PA (WOES/) WOESNER J P.  
PA (HAAS/) HAAS W D.  
PA (GARC/) GARCIA C A.  
PA (KRICK/) KRICKER M.  
PA (SLAT/) SLATER T.  
PA (DAVI/) DAVIS K R.  
PA (ALLE/) ALLEN K.  
PA (HOFF/) HOFFMAN N.  
PA (HURE/) HUREBAN P.  
XX  
PI Gorlach J, An Y, Hamilton CM, Price JL, Raines TM, Yu Y,  
PI Rameeka JG, Page A, Mathew AV, Ledford BL, Woessner JP, Haas WD;  
PI Garcia CA, Kricker M, Slater T, Davis KR, Allen K, Hoffman N;  
PI Hurban P;  
XX  
XX WPI; 2002-403163/43.  
PS  
CC Claim 1; SEQ ID NO 88; 49pp + Sequence Listing; English.  
CC  
CC The invention relates to an Arabidopsis thaliana nucleic acid (I)  
CC comprising a sequence capable of hybridizing under stringent conditions  
CC to a sequence selected from any one of 999 sequences (ABN98233-ABN99231),  
CC given in the specification or its fragment. A polypeptide (II) encoded by  
CC (I), a transgenic plant (III) comprising an exogenous nucleic acid, or a  
CC genetically modified cell (IV) comprising an exogenous nucleic acid, is  
CC useful for screening a candidate agent for its biological effect. (I) is  
CC useful in identifying homologous or related genes, in producing  
CC compositions that modulate the expression or function of its encoded  
CC protein, mapping functional regions of the protein and in studying  
CC associated physiological pathways. (I) is also useful for the genetic  
CC manipulation of cells, particularly plant cells. (I) is also useful in  
CC screening assays of various plant strains to determine the strains that  
CC are best capable of withstanding a particular disease or environmental  
CC stress. (II) and (III) are useful for screening of biologically active  
CC agents, e.g. fungicides, insecticides, etc., for elucidating biochemical  
CC pathways. The screened agents are useful in improved methods of treating  
CC crops to prevent or treat disease. (II) are also useful in screening  
CC programs to identify agents that mimic or enhance the action of tolerance  
CC factors. Such agents are useful in improved methods of treating crops to  
CC enhance their tolerance to environmental stress. (I) is also useful for  
CC enhancing or inhibiting production of a biosynthetic product in a plant.  
CC (III) is useful for identifying other mediators that may induce  
CC expression of proteins of interest, for establishing the extent to which  
CC any specific insect and/or pathogen is responsible for damage to a  
CC particular plant, for identifying other mediators that enhance or induce

Query Match	Best Local Similarity	Score	DB	Length
Matches 337; Conservative 0; Mismatch 284; Indels 3; Gaps 1;	9.8%; 54.0%; Pred. No. 3,4e-35;	153.6;	1124;	
Sequence 1124 BP; 276 A; 267 C; 228 G; 347 T; 0 U; 6 Other;				
CC biosynthetic pathways of nutritional, commercial or medicinal value and				
CC for identifying productions of nutritional, commercial or medicinal				
CC value, (IV) is useful in the study of genetic function and regulation,				
CC for alteration of the cellular metabolism and for screening compounds				
CC Note: The sequence data for this patent did not form part of the printed				
CC specification, but was obtained in electronic format directly from USPTO				
CC at segdata.uspto.gov/sequence.html?docid=593909770445				
XX				
CC				
SO				
Qy 914 AGCGCAAGATGTGACGGTTGAGACATTTGTGAGAGAGTGTAAAGCTTTTCTTCGCG 973				
Db 667 AACGAAACGATGATGATCCGAGATCTGATGAGAGAGTGAAGTTGTTCTATTTCGCG 608				
Qy 974 GGAAGAGACACTTCTATCTGTGACGCGGACGACACTCTGCTATCATGACACCGCG 1053				
Db 607 GGCAGAGACACATCATGACTCTTGTGTTGGCAATGGTCTCTTTAATCAACACGAG 548				
Qy 1034 AGTGCAGGCCCAAGACAGTGTAGGTCTCTCAGGCTGCGCGCTCACGATGTCCCTA 1093				
Db 547 ATTGGCAGGCTGTGTACAGAAAGATGAGCAAGTGTGGC--GATAAAGAACTGG 491				
Qy 1094 CCAAGGACCATGCTTGAAGCTTAAACGTGTAGTGTATGATCTGAACGAGCTTTAAGGT 1153				
Db 490 ATGCAAGAGATCTCAACGACTCAAAATTGTAGCAATGATTAATATGAGTCTTAAAGC 431				
Qy 1154 TGTATCCACCAATAGTAGTCTACGATTGACGCGCTTAATCGATGTAAAGCTAGAGGCT 1213				
Db 430 TATATCTTCATATCCTTCACTGACGTGACAGACCAATTCACAAAGATGAGGCTGTGAGATC 371				
Qy 1214 ACAAAATCCCATGTGGACGAGGCTTATATCCCATCAATPACGGGTCATCATGACCAAG 1273				
Db 370 TGACACTACAGGCGCGCTTGTATCAATCTACTCTTCTTAAGTCCAAACGCAACGG 311				
Qy 1274 CCAATTGGGGTATATGACGTGAACGAATTCATCACTGCGTTTGCGATGAGATGCCGC 1353				
Db 310 AGCTGTGGGGAACCATGACAGAGATTCAGCAAGAGATTCAAAGCGTCTCTTAA 251				
Qy 1334 GTGCTGCCAAACCCGTTGGCTTACCGTTTGACCTCGAGAGTGTGTATGACATGCAATG 1393				
Db 250 AGGCAACAAAGAACCAACCTCTCTTTCTTTGCGGGGACCGAGAAATCGCAATG 191				
Qy 1394 GTGAGATTTTCTATATTGAGGCCAAATGACATCGCTGTATGATCCAGCTTCA 1453				
Db 190 GCGAGAAATTTGATCTGTGAGGCAAGATGCAATGCAATGAACTCCAGAAATCT 131				
Qy 1454 CCTTTCACCTGGCTCTCACTTATAGCATGCAACTACGTCCTTATGTTGCTTATCTCT 1513				
Db 130 CCTTGAGTTTCTCTTCTATGTTTACGCACTTACACAGCTTCCACATTCACCAC 71				
Qy 1514 AACATGTGACCAATCACTCTTC 1537				
Db 70 AGTTGCGTCTCTCTTATCATGC 47				
RESULT 13				
ID ADA48675				
ID ADA48675 standard; DNA; 1377 BP.				
XX ADA48675;				
XX 20-NOV-2003 (first entry)				
XX Rice gene conferring disease resistance in plants.				
XX disease resistance; pathogen tolerance; plant pathogen; ds; gene; plant.				
XX				



XX WO2003000906-A2.  
 XX 03-JAN-2003.  
 XX 21-JUN-2002; 2002WO-IB002453.  
 XX 22-JUN-2001; 2001US-0300112P.  
 XX 26-SEP-2001; 2001US-0352277P.  
 XX 22-MAR-2002; 2002US-036535P.  
 XX (SYGN) SYNGENTA PARTICIPATIONS AG.  
 XX Glaabrook J, Briggs S, Cooper B, Goff SA, Moughamer T;  
 XX Katsgiri F, Kreps J, Provart N, Rieke D, Zhu T;  
 XX WPI; 2003-184052/18.  
 XX P-Psdb; ADA48676.  
 XX New polynucleotide comprising a plant nucleotide sequence having an open  
 XX reading frame that encodes a polypeptide associated with disease  
 XX resistance, useful for conferring resistance or tolerance to a plant  
 XX pathogen.  
 XX Claim 1; SEQ ID NO 745; 299pp; English.  
 XX The invention relates to a novel isolated polynucleotide comprising a  
 XX plant nucleotide sequence having an open reading frame that encodes a  
 XX polypeptide associated with disease resistance or its fragment having  
 XX substantially the same activity as the full-length polypeptide. The  
 XX polynucleotide of the invention is useful for conferring resistance or  
 XX tolerance to a plant pathogen. The present sequence represents a gene  
 XX conferring disease resistance used in the invention.  
 XX Sequence 1377 BP; 324 A; 358 C; 389 G; 306 T; 0 U; 0 Other;  
 XX  
 XX Query Match 9.1%; Score 141.8; DB 8; Length 1377;  
 XX Best Local Similarity 54.1%; Pred. No. 1,4e-31;  
 XX Matches 312; Conservative 0; Mismatches 262; Indels 3; Gaps 1;  
 XX  
 XX 970 GCCGGGAAACAGACACTTCTTAATCTGTCAGCGGACGACCATCTTCTATCCATGAC 1029  
 XX 796 GCATGGGATGACACATCATCACTGTCACCTGGACCATGTCCTGTCAGACGAC 855  
 XX 1030 CCGAGTGGGAGGCGAAGACGTCGATGAGTCTTCAGGCTGCGGCTCAGCGATGTC 1089  
 XX 856 CCAGATTGGGAGGAAATTAAAGGAGAGATTGCAATGAGTGTGTGAC--AGGTG 912  
 XX 1090 CCTACCAAGACATGTCCTTAAGCTTAAACGTTGAGTATGATCTTGAAGAGTCTTTA 1149  
 XX 913 CCACCTGATGACATGCTCAACAGCTAAAGATGTCACATGTCCTCTAGACCCCTA 972  
 XX 1150 AGGTGATCCACCAAGTATGATGATGATTCGACGCGGCTTAATGCGATGTAAGTACGA 1209  
 XX 973 AGGCTGTACAGCCCTGCTTCACTACATACGAGAGAGTTGACATGATGAGTGGC 1032  
 XX 1210 GGGTACAAATCCCATGTCGACGAGCTTCTAATCCCATGATAGCGGCTCATGAC 1269  
 XX 1033 GGCACTAAGATGCTGAGGCTGACATTAATGACCATCCGATGACCAATTCACCGGAC 1092  
 XX 1270 CAAGCATTTGGGGTATGACGTCGATGCAATTAATCCAGCTCGGTTGGGAGTGGAGTG 1329  
 XX 1093 AAGGAATGTGGGGGAGATGCGGATGATTTAGCCGGAAGTTGAAATGGGGTG 1152  
 XX 1330 CCGGTGCTGCAAAACACCCGTTGCTTATACGCTTGTGAGTTCGTACATGC 1389  
 XX 1153 ACGAGGCGAGGAGGAGCCCAACGCACTACTCTCTTCCAGCGGACCGAGGCTATGC 1212  
 XX 1390 ATTGGTCAGATCTTGTATATCTTACGCGCAATTAACACTGCTGTATATCCACGCG 1449  
 XX 1213 ATTGGGAGAACTTTGATGATGATGAGCCCAAGGCTGTATGCTATATCTTACAGAG 1272  
 XX 1450 TTACCTTTCACTTGCTCTTACTTATACAGATGACCTTACGCTTATGTTGCTTAT 1509

DB 1273 TTCTCCTTCACTTGTGCTCCCAAGTACGTCATGACACCCATGATGATGATGATGCTGCT 1332  
 QY 1510 CCTACACATGCTGACCAATCATCACTTCCGAGATTGA 1546  
 DB 1333 CCAAGTACGGGCTCTCTATGATCTTCAAGAGCTCA 1369  
 RESULT 14  
 ID AAC54724/C  
 AAC54724 standard; DNA; 456 BP.  
 AC AAC54724;  
 XX 18-OCT-2000 (first entry)  
 XX Arabidopsis thaliana DNA fragment SEQ ID NO: 78857.  
 XX Hybridisation assay; genetic mapping; gene expression control;  
 XX protein identification; signal transduction pathway; metabolic pathway;  
 XX promoter; termination sequence; ss.  
 XX Arabidopsis thaliana.  
 XX EP1033405-A2.  
 XX 06-SEP-2000.  
 PD 25-FEB-2000; 2000EP-00301439.  
 PF 25-FEB-1999; 99US-0121825P.  
 XX 05-MAR-1999; 99US-0123180P.  
 XX 09-MAR-1999; 99US-0123548P.  
 XX 23-MAR-1999; 99US-0125788P.  
 XX 25-MAR-1999; 99US-0126264P.  
 XX 29-MAR-1999; 99US-0126785P.  
 XX 01-APR-1999; 99US-0127462P.  
 XX 06-APR-1999; 99US-0128234P.  
 XX 08-APR-1999; 99US-0128714P.  
 XX 16-APR-1999; 99US-0129845P.  
 XX 19-APR-1999; 99US-0130077P.  
 XX 21-APR-1999; 99US-0130449P.  
 XX 23-APR-1999; 99US-0130510P.  
 XX 23-APR-1999; 99US-0130891P.  
 XX 28-APR-1999; 99US-0131449P.  
 XX 30-APR-1999; 99US-0132048P.  
 XX 30-APR-1999; 99US-0132477P.  
 XX 04-MAY-1999; 99US-0132484P.  
 XX 05-MAY-1999; 99US-0132485P.  
 XX 06-MAY-1999; 99US-0132486P.  
 XX 06-MAY-1999; 99US-0132487P.  
 XX 07-MAY-1999; 99US-0132863P.  
 XX 11-MAY-1999; 99US-0134265P.  
 XX 14-MAY-1999; 99US-0134218P.  
 XX 14-MAY-1999; 99US-0134219P.  
 XX 14-MAY-1999; 99US-0134221P.  
 XX 14-MAY-1999; 99US-0134370P.  
 XX 18-MAY-1999; 99US-0134768P.  
 XX 19-MAY-1999; 99US-0134941P.  
 XX 20-MAY-1999; 99US-0135124P.  
 XX 21-MAY-1999; 99US-0135353P.  
 XX 24-MAY-1999; 99US-0135629P.  
 XX 25-MAY-1999; 99US-0136031P.  
 XX 27-MAY-1999; 99US-0136382P.  
 XX 28-MAY-1999; 99US-0136782P.  
 XX 01-JUN-1999; 99US-0137222P.  
 XX 03-JUN-1999; 99US-0137528P.  
 XX 04-JUN-1999; 99US-0137502P.  
 XX 07-JUN-1999; 99US-0137724P.  
 XX 08-JUN-1999; 99US-0138094P.  
 XX 10-JUN-1999; 99US-0138540P.  
 XX 10-JUN-1999; 99US-0138647P.  
 XX 14-JUN-1999; 99US-0139119P.

PR 16-JUN-1999; 99US-0139452P.  
 PR 16-JUN-1999; 99US-0139453P.  
 PR 17-JUN-1999; 99US-0139454P.  
 PR 18-JUN-1999; 99US-0139455P.  
 PR 18-JUN-1999; 99US-0139456P.  
 PR 18-JUN-1999; 99US-0139457P.  
 PR 18-JUN-1999; 99US-0139458P.  
 PR 18-JUN-1999; 99US-0139459P.  
 PR 18-JUN-1999; 99US-0139460P.  
 PR 18-JUN-1999; 99US-0139461P.  
 PR 18-JUN-1999; 99US-0139462P.  
 PR 18-JUN-1999; 99US-0139463P.  
 PR 18-JUN-1999; 99US-0139750P.  
 PR 18-JUN-1999; 99US-0139753P.  
 PR 21-JUN-1999; 99US-0139817P.  
 PR 22-JUN-1999; 99US-0139899P.  
 PR 23-JUN-1999; 99US-0140353P.  
 PR 23-JUN-1999; 99US-0140354P.  
 PR 24-JUN-1999; 99US-0140695P.  
 PR 28-JUN-1999; 99US-0140823P.  
 PR 29-JUN-1999; 99US-0140991P.  
 PR 30-JUN-1999; 99US-0141287P.  
 PR 01-JUL-1999; 99US-0141842P.  
 PR 01-JUL-1999; 99US-0142154P.  
 PR 02-JUL-1999; 99US-0142055P.  
 PR 06-JUL-1999; 99US-0142390P.  
 PR 08-JUL-1999; 99US-0142803P.  
 PR 09-JUL-1999; 99US-0142920P.  
 PR 12-JUL-1999; 99US-0142977P.  
 PR 13-JUL-1999; 99US-0143542P.  
 PR 14-JUL-1999; 99US-0143624P.  
 PR 15-JUL-1999; 99US-0144005P.  
 PR 16-JUL-1999; 99US-0144086P.  
 PR 16-JUL-1999; 99US-0144086P.  
 PR 19-JUL-1999; 99US-0144332P.  
 PR 19-JUL-1999; 99US-0144332P.  
 PR 19-JUL-1999; 99US-0144332P.  
 PR 19-JUL-1999; 99US-0144333P.  
 PR 19-JUL-1999; 99US-0144333P.  
 PR 19-JUL-1999; 99US-0144334P.  
 PR 20-JUL-1999; 99US-0144335P.  
 PR 20-JUL-1999; 99US-0144632P.  
 PR 20-JUL-1999; 99US-0144804P.  
 PR 21-JUL-1999; 99US-0144814P.  
 PR 21-JUL-1999; 99US-0145088P.  
 PR 21-JUL-1999; 99US-0145088P.  
 PR 22-JUL-1999; 99US-0145085P.  
 PR 22-JUL-1999; 99US-0145087P.  
 PR 22-JUL-1999; 99US-0145089P.  
 PR 22-JUL-1999; 99US-0145192P.  
 PR 23-JUL-1999; 99US-0145192P.  
 PR 23-JUL-1999; 99US-0145145P.  
 PR 23-JUL-1999; 99US-0145218P.  
 PR 23-JUL-1999; 99US-0145224P.  
 PR 26-JUL-1999; 99US-0145276P.  
 PR 27-JUL-1999; 99US-0145913P.  
 PR 27-JUL-1999; 99US-0145918P.  
 PR 27-JUL-1999; 99US-0145919P.  
 PR 28-JUL-1999; 99US-0145951P.  
 PR 28-JUL-1999; 99US-0146386P.  
 PR 02-AUG-1999; 99US-0146386P.  
 PR 02-AUG-1999; 99US-0146388P.  
 PR 02-AUG-1999; 99US-0146389P.  
 PR 03-AUG-1999; 99US-0147038P.  
 PR 04-AUG-1999; 99US-0147040P.  
 PR 04-AUG-1999; 99US-0147022P.  
 PR 05-AUG-1999; 99US-0147192P.  
 PR 05-AUG-1999; 99US-0147260P.  
 PR 06-AUG-1999; 99US-0147303P.  
 PR 06-AUG-1999; 99US-0147416P.  
 PR 06-AUG-1999; 99US-0147439P.  
 PR 09-AUG-1999; 99US-0147935P.  
 PR 09-AUG-1999; 99US-0148171P.  
 PR 10-AUG-1999; 99US-0148319P.  
 PR 11-AUG-1999; 99US-0148319P.

PR 12-AUG-1999; 99US-0148341P.  
 PR 13-AUG-1999; 99US-0148565P.  
 PR 13-AUG-1999; 99US-0148684P.  
 PR 16-AUG-1999; 99US-0149368P.  
 PR 17-AUG-1999; 99US-0149175P.  
 PR 18-AUG-1999; 99US-0149426P.  
 PR 20-AUG-1999; 99US-0149722P.  
 PR 20-AUG-1999; 99US-0149723P.  
 PR 20-AUG-1999; 99US-0149929P.  
 PR 23-AUG-1999; 99US-0149930P.  
 PR 25-AUG-1999; 99US-0150566P.  
 PR 26-AUG-1999; 99US-0150884P.  
 PR 27-AUG-1999; 99US-0151065P.  
 PR 27-AUG-1999; 99US-0151066P.  
 PR 27-AUG-1999; 99US-0151080P.  
 PR 30-AUG-1999; 99US-0151303P.  
 PR 31-AUG-1999; 99US-0151438P.  
 PR 01-SEP-1999; 99US-0151930P.  
 PR 07-SEP-1999; 99US-0152363P.  
 PR 10-SEP-1999; 99US-0153070P.  
 PR 13-SEP-1999; 99US-0153758P.  
 PR 15-SEP-1999; 99US-0154018P.  
 PR 16-SEP-1999; 99US-0154039P.  
 PR 20-SEP-1999; 99US-0154779P.  
 PR 22-SEP-1999; 99US-0155139P.  
 PR 23-SEP-1999; 99US-0155486P.  
 PR 24-SEP-1999; 99US-0155659P.  
 PR 28-SEP-1999; 99US-0156458P.  
 PR 29-SEP-1999; 99US-0156536P.  
 PR 04-OCT-1999; 99US-0157117P.  
 PR 05-OCT-1999; 99US-0157753P.  
 PR 06-OCT-1999; 99US-0157665P.  
 PR 07-OCT-1999; 99US-0158023P.  
 PR 08-OCT-1999; 99US-0158222P.  
 PR 12-OCT-1999; 99US-0158369P.  
 PR 13-OCT-1999; 99US-0159293P.  
 PR 13-OCT-1999; 99US-0159294P.  
 PR 13-OCT-1999; 99US-0159295P.  
 PR 14-OCT-1999; 99US-0159329P.  
 PR 14-OCT-1999; 99US-0159333P.  
 PR 14-OCT-1999; 99US-0159637P.  
 PR 14-OCT-1999; 99US-0159638P.  
 PR 18-OCT-1999; 99US-0159584P.  
 PR 21-OCT-1999; 99US-0160741P.  
 PR 21-OCT-1999; 99US-0160767P.  
 PR 21-OCT-1999; 99US-0160768P.  
 PR 21-OCT-1999; 99US-0160770P.  
 PR 21-OCT-1999; 99US-0160815P.  
 PR 22-OCT-1999; 99US-0160815P.  
 PR 22-OCT-1999; 99US-0160980P.  
 PR 22-OCT-1999; 99US-0160981P.  
 PR 22-OCT-1999; 99US-0161404P.  
 PR 25-OCT-1999; 99US-0161405P.  
 PR 25-OCT-1999; 99US-0161406P.  
 PR 26-OCT-1999; 99US-0161359P.  
 PR 26-OCT-1999; 99US-0161360P.  
 PR 26-OCT-1999; 99US-0161361P.  
 PR 28-OCT-1999; 99US-0161920P.  
 PR 28-OCT-1999; 99US-0161923P.  
 PR 28-OCT-1999; 99US-0161935P.  
 PR 29-OCT-1999; 99US-0162142P.

Query Match 6.7%; Score 104.2; DB 3; Length 456;  
 Best Local Similarity 52.7%; Pred. No. 1.6e-20;  
 Matches 226; Conservative 0; Mismatches 203; Indels 0; Gaps 0;

QY 1109 TTAAGCTTAACGTTGACGTGATGATCTTGAACGAGCTCTTTAAGGTGATCCACCAATAG 1168  
 Db 439 TTTGTCTCCTACGATGACCAAGATTATATATGAGCTCTGAGGCTATATCTCTCCAGTAA 380

Mon May 24 08:10:07 2004

us-09-992-901-1.rng

Page 18

```
QY      1169  TAGCTACGATTCGACGCGCTAAATCGATGTGAGAGGATACAAAATCCCATGTG 1228
DB      379  TCCAGTGAACAGAGGATCTCAAAAGAGATTAAGCTAGGTATGTACACTACAGCGC 320
QY      1229  GACGAGGCTTCTATCCCATTCATGCGGCTCCATCATGACCAAGCCATTGGGGTATG 1288
DB      319  GATATCAAGTCCATATGCGCTGTTCTGCTAATCCACCGACACTAAGCTTGGGGCGATG 260
QY      1289  ACGTGAACGATTCATTCACGATCGGTTGGCCGATGAGATGGCGCGCTGCTCCAAACACC 1348
DB      259  ATCCAGCGGAGTTCAAGCCCGAGAGGTTCAAAGACGGATCCGCAAGGCAACAAAGAAC 200
QY      1349  CCGTTGGCTTCATACCGCTTGGCTTCGAGTTGCTGATGATGATGATGATGATGATGATG 1408
DB      199  AGGTCGCTCTCTGCTTCCTTCGGGTGGGACCGAGGATTTGCATCGTCAAGACTTCGCTC 140
QY      1409  TACTTACGCGCAATTCGACACTCGCTGTATGATGATGATGATGATGATGATGATGATG 1468
DB      139  TGTGGAGCAAGATGAGCTGTGGCTTTATTTCTTCAGAGATTCTCCCTTCAGACTTTCAC 80
QY      1469  CTAATTATCAGATGACCTACCGTCCCTTATGCTTATGCTTATGCTTATGCTTATGCTTAT 1528
DB      79  CTTCCTATGTTTCAATTCACCTTACAGTCTTCACTATACATCATCCACAGTGTGGGCTCATC 20
QY      1529  TCACCTTCC 1537
DB      19  TTATCTCTG 11

RESULT 15
AAC53115/c
ID      AAC53115 standard; DNA; 458 BP.
XX
AC      AAC53115;
XX
DT      18-OCT-2000 (first entry)
XX
DE      Arabidopsis thaliana DNA fragment SEQ ID NO: 73436.
XX
KM      Hybridisation assay; genetic mapping; gene expression control;
KW      protein identification; signal transduction pathway; metabolic pathway;
XX      promoter; termination sequence; ss.
XX
OS      Arabidopsis thaliana.
XX
FN      EP1033405-A2.
XX
PD      06-SEP-2000.
XX
PF      25-FEB-2000; 2000EP-00301439.
XX
XX      25-FEB-1999; 99US-0121825P.
PR      05-MAR-1999; 99US-0133180P.
PR      09-MAR-1999; 99US-0133588P.
PR      23-MAR-1999; 99US-0125788P.
PR      25-MAR-1999; 99US-0126264P.
PR      29-MAR-1999; 99US-0126785P.
PR      01-APR-1999; 99US-0127462P.
PR      06-APR-1999; 99US-0128234P.
PR      08-APR-1999; 99US-0128714P.
PR      16-APR-1999; 99US-0129845P.
PR      19-APR-1999; 99US-0130077P.
PR      21-APR-1999; 99US-0130449P.
PR      23-APR-1999; 99US-0130510P.
PR      28-APR-1999; 99US-0130891P.
PR      30-APR-1999; 99US-0132048P.
PR      30-APR-1999; 99US-0132407P.
PR      04-MAY-1999; 99US-0132484P.
PR      05-MAY-1999; 99US-0132485P.
PR      06-MAY-1999; 99US-0132486P.
PR      06-MAY-1999; 99US-0132487P.
PR      07-MAY-1999; 99US-0132863P.
PR      11-MAY-1999; 99US-0134266P.
PR      14-MAY-1999; 99US-0134218P.
PR      14-MAY-1999; 99US-0134219P.
PR      14-MAY-1999; 99US-0134221P.
PR      14-MAY-1999; 99US-0134370P.
PR      18-MAY-1999; 99US-0134768P.
PR      19-MAY-1999; 99US-0134941P.
PR      20-MAY-1999; 99US-0135124P.
PR      21-MAY-1999; 99US-0135353P.
PR      24-MAY-1999; 99US-0135629P.
PR      25-MAY-1999; 99US-0136021P.
PR      27-MAY-1999; 99US-0136392P.
PR      28-MAY-1999; 99US-0136782P.
PR      01-JUN-1999; 99US-0137223P.
PR      03-JUN-1999; 99US-0137528P.
PR      04-JUN-1999; 99US-0137502P.
PR      07-JUN-1999; 99US-0137724P.
PR      08-JUN-1999; 99US-0138094P.
PR      10-JUN-1999; 99US-0138540P.
PR      10-JUN-1999; 99US-0138847P.
PR      14-JUN-1999; 99US-0139119P.
PR      16-JUN-1999; 99US-0139452P.
PR      16-JUN-1999; 99US-0139453P.
PR      17-JUN-1999; 99US-0139492P.
PR      17-JUN-1999; 99US-0139454P.
PR      18-JUN-1999; 99US-0139455P.
PR      18-JUN-1999; 99US-0139456P.
PR      18-JUN-1999; 99US-0139457P.
PR      18-JUN-1999; 99US-0139458P.
PR      18-JUN-1999; 99US-0139459P.
PR      18-JUN-1999; 99US-0139460P.
PR      18-JUN-1999; 99US-0139461P.
PR      18-JUN-1999; 99US-0139462P.
PR      18-JUN-1999; 99US-0139463P.
PR      18-JUN-1999; 99US-0139750P.
PR      18-JUN-1999; 99US-0139763P.
PR      21-JUN-1999; 99US-0139817P.
PR      22-JUN-1999; 99US-0139899P.
PR      23-JUN-1999; 99US-0140353P.
PR      23-JUN-1999; 99US-0140354P.
PR      24-JUN-1999; 99US-0140695P.
PR      26-JUN-1999; 99US-0140823P.
PR      29-JUN-1999; 99US-0140891P.
PR      30-JUN-1999; 99US-0141287P.
PR      01-JUL-1999; 99US-0141542P.
PR      02-JUL-1999; 99US-0142055P.
PR      06-JUL-1999; 99US-0142390P.
PR      08-JUL-1999; 99US-0142803P.
PR      09-JUL-1999; 99US-0142920P.
PR      12-JUL-1999; 99US-0142977P.
PR      13-JUL-1999; 99US-0143542P.
PR      14-JUL-1999; 99US-0143624P.
PR      15-JUL-1999; 99US-0144005P.
PR      16-JUL-1999; 99US-0144085P.
PR      16-JUL-1999; 99US-0144086P.
PR      19-JUL-1999; 99US-0144325P.
PR      19-JUL-1999; 99US-0144331P.
PR      19-JUL-1999; 99US-0144332P.
PR      19-JUL-1999; 99US-0144333P.
PR      19-JUL-1999; 99US-0144334P.
PR      19-JUL-1999; 99US-0144335P.
PR      20-JUL-1999; 99US-0144532P.
PR      20-JUL-1999; 99US-0144632P.
PR      20-JUL-1999; 99US-0144884P.
PR      21-JUL-1999; 99US-0144814P.
PR      21-JUL-1999; 99US-0145086P.
PR      21-JUL-1999; 99US-0145088P.
PR      22-JUL-1999; 99US-0145085P.
PR      22-JUL-1999; 99US-0145087P.
PR      22-JUL-1999; 99US-0145089P.
PR      22-JUL-1999; 99US-0145192P.
PR      23-JUL-1999; 99US-0145145P.
```

Mon May 24 08:10:07 2004

us-09-992-901-1.rng

Page 19

PR 23-JUL-1999; 99US-0155228P.  
PR 23-JUL-1999; 99US-0155228P.  
PR 26-JUL-1999; 99US-0145226P.  
PR 27-JUL-1999; 99US-0145913P.  
PR 27-JUL-1999; 99US-0145913P.  
PR 27-JUL-1999; 99US-0145913P.  
PR 28-JUL-1999; 99US-0145913P.  
PR 02-AUG-1999; 99US-0145951P.  
PR 02-AUG-1999; 99US-0146386P.  
PR 02-AUG-1999; 99US-0146386P.  
PR 03-AUG-1999; 99US-0147038P.  
PR 04-AUG-1999; 99US-0147204P.  
PR 04-AUG-1999; 99US-0147302P.  
PR 05-AUG-1999; 99US-0147192P.  
PR 05-AUG-1999; 99US-0147260P.  
PR 06-AUG-1999; 99US-0147303P.  
PR 06-AUG-1999; 99US-0147416P.  
PR 09-AUG-1999; 99US-0147935P.  
PR 09-AUG-1999; 99US-0147935P.  
PR 10-AUG-1999; 99US-0148312P.  
PR 11-AUG-1999; 99US-0148312P.  
PR 12-AUG-1999; 99US-0148341P.  
PR 13-AUG-1999; 99US-0148565P.  
PR 13-AUG-1999; 99US-0148684P.  
PR 16-AUG-1999; 99US-0149368P.  
PR 17-AUG-1999; 99US-0149175P.  
PR 18-AUG-1999; 99US-0149422P.  
PR 20-AUG-1999; 99US-0149722P.  
PR 20-AUG-1999; 99US-0149723P.  
PR 20-AUG-1999; 99US-0149929P.  
PR 23-AUG-1999; 99US-0149902P.  
PR 23-AUG-1999; 99US-0149930P.  
PR 25-AUG-1999; 99US-0150566P.  
PR 26-AUG-1999; 99US-0151084P.  
PR 27-AUG-1999; 99US-0151065P.  
PR 27-AUG-1999; 99US-0151066P.  
PR 27-AUG-1999; 99US-0151080P.  
PR 27-AUG-1999; 99US-0151080P.  
PR 30-AUG-1999; 99US-0151303P.  
PR 31-AUG-1999; 99US-0151303P.  
PR 01-SEP-1999; 99US-0151930P.  
PR 07-SEP-1999; 99US-0152363P.  
PR 10-SEP-1999; 99US-0153707P.  
PR 13-SEP-1999; 99US-0153758P.  
PR 15-SEP-1999; 99US-0154018P.  
PR 16-SEP-1999; 99US-0154039P.  
PR 20-SEP-1999; 99US-0154799P.  
PR 22-SEP-1999; 99US-0155139P.  
PR 23-SEP-1999; 99US-0155486P.  
PR 24-SEP-1999; 99US-0155659P.  
PR 28-SEP-1999; 99US-0156458P.  
PR 29-SEP-1999; 99US-0156596P.  
PR 04-OCT-1999; 99US-0157117P.  
PR 05-OCT-1999; 99US-0157753P.  
PR 06-OCT-1999; 99US-0157865P.  
PR 07-OCT-1999; 99US-0158029P.  
PR 08-OCT-1999; 99US-0158232P.  
PR 12-OCT-1999; 99US-0158369P.  
PR 13-OCT-1999; 99US-0159293P.  
PR 13-OCT-1999; 99US-0159294P.  
PR 14-OCT-1999; 99US-0159329P.  
PR 14-OCT-1999; 99US-0159330P.  
PR 14-OCT-1999; 99US-0159331P.  
PR 14-OCT-1999; 99US-0159337P.  
PR 14-OCT-1999; 99US-0159637P.  
PR 18-OCT-1999; 99US-0159684P.  
PR 21-OCT-1999; 99US-0160741P.  
PR 21-OCT-1999; 99US-0160767P.  
PR 21-OCT-1999; 99US-0160768P.  
PR 21-OCT-1999; 99US-0160770P.  
PR 21-OCT-1999; 99US-0160814P.  
PR 21-OCT-1999; 99US-0160815P.  
PR 22-OCT-1999; 99US-0160980P.

PR 22-OCT-1999; 99US-0160981P.  
PR 22-OCT-1999; 99US-0160989P.  
PR 25-OCT-1999; 99US-0161404P.  
PR 25-OCT-1999; 99US-0161405P.  
PR 25-OCT-1999; 99US-0161405P.  
PR 25-OCT-1999; 99US-0161406P.  
PR 26-OCT-1999; 99US-0161359P.  
PR 26-OCT-1999; 99US-0161360P.  
PR 26-OCT-1999; 99US-0161361P.  
PR 28-OCT-1999; 99US-0161920P.  
PR 28-OCT-1999; 99US-0161920P.  
PR 28-OCT-1999; 99US-0161923P.  
PR 29-OCT-1999; 99US-0162142P.

Query Match 6.6%; Score 102.6; DB 3; Length 458;  
Best Local Similarity 52.4%; Pred. No. 4; ee-20;  
Matches 225; Conservative 0; Mismatches 204; Indels 0; Gaps 0;

QY 1109 TTAAGCTTAAACCTTGAATGATATCTTGAACGAGCTTTAAGTTGATCCACCAATAG 1168  
DB 441 TTTGTCTTCAATGACATGATTTATATATGAGTCTGAGCTATATCTCCAGTAA 382  
QY 1169 TAGCTACGATTCGACGCGCTAAATCGGATGTGAAGCTAGGAGGCTACAAATCCCATGTG 1228  
DB 381 TCCAGATGACAGGCGGCTCACAAGAGATTAAGCTAGGTGATATGACACTACAGGCT 322  
QY 1229 GACGAGCTTCTATATCCCATCTATAGCGCTCCATCTATGACCAAGCCATTTGGGGTAATG 1288  
DB 321 GTATTCAAGTCCATATGCTGTCTGTCTAAATCCACCGTGAOACTAAGCTTTGGGGCGATG 262  
QY 1289 ACGTGAACGAATTCATCCAGCTCGCTTGGGATGAGTGGCGGTCTGCCAACC 1348  
DB 261 ATGACGCGGAGTTCAAGCCGAGAGCTTCAAGAGGATCGCAAGGCAACAAAGACC 202  
QY 1349 CCGTGGCTTCAATCCGTTTGGCTCGAGTTCTATATGATGATTTGGTCAAGATCTTGCTA 1408  
DB 201 AGCTCTGCTTCTGCTTCGCGTGGGACCGAGATTTGATCGTCAAGTCTTCCCTC 142  
QY 1409 TACTTCAGGCAAAATGACACTCGCTGATATGATTCACAGCTTCACTTGTGCTC 1468  
DB 141 TATTGAAAGCAAAATGAGCTCTGCTTGTATCTTCAGAGATTCCTTGAAGCTTTCAC 82  
QY 1469 CTACTTATCAGATCAGCTTACCGCTCTTATGTTGCTTATCTTCACATGATGACCAA 1528  
DB 81 CTTCCTATGATTCATTCACCTTACAGAGCTTCACTATATACATCCACAGTGTGGGGCTATC 22  
QY 1529 TCACCTTGC 1537  
DB 21 TTATCCTGC 13

Search completed: May 22, 2004, 17:15:54  
Job time : 695 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 22, 2004, 16:11:26 ; Search time 4335 Seconds  
(without alignments)

10766.918 Million cell updates/sec

Title: US-09-992-901-1  
Perfect score: 1563  
Sequence: 1 atgagagaagaagtagcag.....tgacacatcatgagattga 1563

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapekt 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : EST.\*  
1: em\_estb1.\*  
2: em\_estb2.\*  
3: em\_estb3.\*  
4: em\_estb4.\*  
5: em\_estb5.\*  
6: em\_estb6.\*  
7: em\_estb7.\*  
8: em\_estb8.\*  
9: em\_estb9.\*  
10: em\_estb10.\*  
11: em\_estb11.\*  
12: em\_estb12.\*  
13: em\_estb13.\*  
14: em\_estb14.\*  
15: em\_estb15.\*  
16: em\_estb16.\*  
17: em\_estb17.\*  
18: em\_estb18.\*  
19: em\_estb19.\*  
20: em\_estb20.\*  
21: em\_estb21.\*  
22: em\_estb22.\*  
23: em\_estb23.\*  
24: em\_estb24.\*  
25: em\_estb25.\*  
26: em\_estb26.\*  
27: em\_estb27.\*  
28: em\_estb28.\*  
29: em\_estb29.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	442	28.3	562	B29839	B29839 T17P14TFC T
2	345.8	22.1	727	B1926826	B1926826 EST546715
3	316.2	20.2	558	B28185	B28185 T4A14TFC TA
4	311.6	19.9	916	CK262340	CK262340 EST708418

Result	Score	Query Match	Length	ID	Description
5	307.8	19.7	614	B0891384	B0891384 P049G01 P
6	287.2	18.4	1876	A1109087	A1109087 Zea mays
7	280.2	17.9	698	CG027129	CG027129 CHGA11TFC
8	277	17.7	807	CK262341	CK262341 EST708419
9	259.6	16.6	817	BH597310	BH597310 BOGT577R
10	256	16.4	493	B0297122	B0297122 sac38e02
11	244	15.6	303	B62694	B62694 T17P24TFC.1
12	242.4	15.5	834	CG112869	CG112869 PUFK51TB
13	234	15.0	1567	A1104671	A1104671 Zea mays
14	232.8	14.9	424	AV417119	AV417119 AV417119
15	219.6	14.0	815	CG336779	CG336779 OGUAE25TV
16	206.2	13.2	377	AU287860	AU287860 AU287860
17	205.4	13.1	511	BJ558923	BJ558923 BJ558923
18	201.4	12.9	675	CP432662	CP432662 NIT1.18 B
19	199.8	12.8	445	BM144017	BM144017 sac35d11
20	198.8	12.7	861	A1107534	A1107534 Zea mays
21	197.6	12.6	533	BG556624	BG556624 EM1.37.E1
22	196.8	12.6	660	BH878919	BH878919 hsb5b04.9
23	196.4	12.6	412	BE210430	BE210430 sc043f12.Y
24	195.2	12.5	474	BG135565	BG135565 EST168457
25	193.2	12.4	730	CG027138	CG027138 CHGA11TFC
26	192.6	12.3	622	AM738972	AM738972 EST39499
27	189.6	12.1	968	BG321478	BG321478 Ds01.07G1
28	188.2	12.0	534	BG316378	BG316378 sabb0402
29	186.6	11.9	550	CD055580	CD055580 H009K04S
30	185.8	11.9	439	BH723464	BH723464 BOMIR22TF
31	185.6	11.9	649	AM622085	AM622085 EST112883
32	184.6	11.8	774	CF672445	CF672445 RTCTWT1.63
33	184	11.8	817	CK253045	CK253045 EST736682
34	183.4	11.7	751	AV920529	AV920529 AV920529
35	182.4	11.7	831	BE705698	BE705698 SC02.01h0
36	180.6	11.6	1055	CK162569	CK162569 FGAS01516
37	180.6	11.6	682	BQ801263	BQ801263 WHE2812.C
38	180	11.5	223	CC054934	CC054934 SALK_0803
39	179.8	11.5	690	AV944488	AV944488 AV944488
40	177.2	11.3	347	BM500512	BM500512 PAC000000
41	176.4	11.3	738	B1925645	B1925645 EST545534
42	176	11.3	670	BG452638	BG452638 NF078A02L
43	175.8	11.2	606	CA901199	CA901199 PCS03751
44	174.6	11.2	715	CD482646	CD482646 atv01-50m
45	174.6	11.2	754	BE574891	BE574891 w_A2R T11

## ALIGNMENTS

RESULT 1  
LOCUS B29839 562 bp DNA linear GSS 13-OCT-1997  
DEFINITION T17P14TFC TAMU Arabidopsis thaliana genomic clone T17P14, genomic survey sequence.  
ACCESSION B29839  
VERSION B29839.1 GI:2515805  
KEYWORDS GSS.  
SOURCE Arabidopsis thaliana (thale cress)  
ORGANISM Arabidopsis thaliana  
Bukharjocai, Viridiplantae, Streptophyta, Embryophyta, Tracheophyta, Spermatophyta, Magnoliophyta, eudicotyledons, core eudicots, rosids, eurosids II, Brassicales, Brassicaceae, Arabidopsids.  
REFERENCE 1 (bases 1 to 562)  
AUTHORS Rounsley, S.D., Kelley, J.M., Field, C.E., Craven, M.B., Adams, M.D. and Venter, J.C.  
TITLE Use of a BAC End Sequence Database To Identify Minimal Overlaps for Arabidopsis Genomic Sequencing  
JOURNAL Unpublished (1997)  
COMMENT Other GSSs: T17P14TFC  
Contact: Steve Rounsley  
Department of Eukaryotic Genomics  
The Institute for Genomic Research  
9712 Medical Center Dr., Rockville, MD 20850, USA  
Tel: 301 838 0208  
Fax: 301 838 0208  
Email: rounsley@igf.org



RESULT 3  
B28185/c  
LOCUS B28185 558 bp DNA linear GSS 13-OCT-1997  
DEFINITION T4A14TFG T4M1 Arabidopsis thaliana genomic clone T4A14, genomic survey sequence.  
ACCESSION B28185  
VERSION B28185.1 GI:2514151  
KEYWORDS GSS.  
SOURCE Arabidopsis thaliana (thale cress)  
ORGANISM Arabidopsis thaliana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.  
REFERENCE 1 (bases 1 to 558)  
AUTHORS Rounsley S.D., Kelley, J.M., Field, C.E., Craven, M.B., Adams, M.D. and Venter, J.C.  
TITLE Use of a PAC End Sequence Database To Identify Minimal Overlaps for Arabidopsis Genomic Sequencing  
JOURNAL Unpublished (1997)  
COMMENT Other GSSs: T4A14TRB  
Contact: Steve Rounsley  
Department of Eukaryotic Genomics  
The Institute for Genomic Research  
9712 Medical Center Dr., Rockville, MD 20850, USA  
Tel: 301 838 0200  
Fax: 301 838 0208  
Email: rounsley@igr.org  
Seq primer: M13-21  
Class: BAC ends  
High quality sequence stop: 558.  
Location/Qualifiers  
1..558  
/organism="Arabidopsis thaliana"  
/mol\_type="genomic DNA"  
/strain="Columbia"  
/db\_xref="taxon:3702"  
/clone="T4A14"  
/sex="hermaphrodite"  
/clone\_1id="T4M1"  
/note="Vector: BelosBAC11; Site\_1: HindIII; Site\_2: HindIII; Produced by Rod Wang"  
ORIGIN  
Query Match 20.2%; Score 316.2; DB 28; Length 558;  
Best Local Similarity 99.1%; Pred. No. 1.1e-79; Indels 0; Gaps 0;  
Matches 318; Conservative 0; Mismatches 3;  
1243 ATCCCAATCATAGCGGTCCATCATGACCAAGCCATTGGGGTAAATGACGTGAACGAATTTC 1302  
Db ATCCCAATCATAGCGGTCCATCATGACCAAGCCATTGGGGTAAATGACGTGAACGAATTTC 499  
Qy 1303 AATCCAGCTCGCTTTGGCGATGAGATGCCGGTGTGTGCGCAACGCCCGTTGGTTCATA 1362  
Db 498 ATTCCAGCTCGCTTTGGCGATGAGATGCCGGTGTGTGCGCAACGCCCGTTGGTTCATA 439  
Qy 1363 CCGTTGGGCTCGAGATTCATCATGATGATGGTGCAGATCTTGCTATATCTTCAGGCCAA 1422  
Db 438 CCGTTGGGCTCGAGATTCATCATGATGATGGTGCAGATCTTGCTATATCTTCAGGCCAA 379  
Qy 1423 TTGAACCTGCTGTAATGATCAACAGCTTACCTTTACCTTGGCTCTTACTATACGAT 1482  
Db 378 TTGAACCTGCTGTAATGATCAACAGCTTACCTTTACCTTGGCTCTTACTATACGAT 319  
Qy 1483 GGACCTACCGCTTATGTTGCTTATCTTCAACATGATGACCAATCACTTCGAGAGA 1542  
Db 318 GGACCTACCGCTTATGTTGCTTATCTTCAACATGATGACCAATCACTTCGAGAGA 259  
Qy 1543 TTGACCAATCATGAGGATTGA 1563  
Db 258 TTGACCAATCATGAGGATTGA 238

RESULT 4  
CK262340  
LOCUS CK262340 916 bp mRNA linear EST 12-DEC-2003  
DEFINITION EST708418 potato abiotic stress cDNA library Solanum tuberosum cDNA clone POABF61 5' end, mRNA sequence.  
ACCESSION CK262340  
VERSION CK262340.1 GI:39819318  
KEYWORDS EST.  
SOURCE Solanum tuberosum (potato)  
ORGANISM Solanum tuberosum  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids; lamids; Solanales; Solanaceae; Solanum.  
REFERENCE 1 (bases 1 to 916)  
AUTHORS Buell, C.R., Hart, A., Zismann, V., Karanycheva, S.A. and Baker, B.  
TITLE Generation of ESTs from abiotic stressed potato tissue  
JOURNAL Unpublished (2003)  
COMMENT Other ESTs: EST708419  
Contact: Robin Buell  
The Institute for Genomic Research  
9712 Medical Center Dr., Rockville, MD 20850, USA  
Email: potato-array@igr.org  
Clones can be requested from TIGR via potato@tigr.org  
Seq primer: ATT TAG GTG ACA CTA TAG.  
Location/Qualifiers  
1..916  
/organism="Solanum tuberosum"  
/mol\_type="mRNA"  
/cultivar="Kennebec"  
/db\_xref="taxon:4113"  
/clone="POABF61"  
/tissue\_type="abiotic stress treated leaf and root tissue"  
/lab\_host="DH10B-Tona"  
/clone\_1id="potato abiotic stress cDNA library"  
/note="Vector: pCMVSPORT6.1; Site\_1: EcoRI; Site\_2: NotI; supplier: Solanum tuberosum var. Kennebec plants were grown from cuttings on a 16hr light/8 hr dark cycle at 25 C for 3-4 weeks. Abiotic stress conditions were applied to four separate sets of plants. Set 1 involved saturation of the soil with 150 mM NaCl and tissues were harvested at following application of the salt stress (leaves: 2hr, 6hr, 12hr, 1d, 2d, and 4d; roots: 2hr, 6hr, 12hr, and 2d). Set 2 were grown under the standard conditions and then were water stressed by withdrawal of further watering applications. Drought stressed plants were harvested after cessation of watering (leaves: 3d, 5d, and 7d; roots: 3d and 5d). Set 3 were grown under the standard conditions and then were cold stressed by placement at 4 C. Cold stressed leaves were harvested at 2 hr, 6 hr, 12 hr, 1 d, and 4d and roots were harvested at 2 hr, 6 hr, 12 hr, 1 d, 2d. Set 4 were grown under the standard conditions and then were heat stressed by placement at 35 C. Heat stressed leaves were harvested at 2 hr, 6 hr, 12 hr, 1 d, 2d and 4d and heat-stressed roots were harvested at 6 hr, 12 hr, 1 d, and 4d. RNA was isolated from all tissues and equal RNA from each tissue and stress was pooled to construct the cDNA library. RNA sample."  
ORIGIN  
Query Match 19.9%; Score 311.6; DB 14; Length 916;  
Best Local Similarity 63.4%; Pred. No. 3e-78;  
Matches 511; Conservative 0; Mismatches 289; Indels 6; Gaps 2;  
37 GTTCTGTTGCTGCTGTAATCTTAACTTATATGTAAGGTATGCTGTTATG 96  
Db 114 GTTATTTGATTTTGTATGTTGTTACTGTTGTTACAGAAAGGCTTTGTTACTTTGG 173  
Qy 97 TGGAGACCAAGAGATTGAAGACATTTCTTAACAGAGATTGAGGCTCTTAT 156  
Db 174 TGGAGACCAAGAGATTGAAGACATTTCTTAACAGAGATTGAGGCTCTTAT 233  
Qy 157 CATTTCTTATGGAATGTTAAAGACTTGTGAATGATGCTTAAAGCTTCTTCAT 216

Db	234	CAATTTCTCTTGGTAACCTGAAGA---ATTGCAAGTTTCACTACTCCATCTTGAGGCT	290
Qy	217	CCTANGCCTTTCTCTCAACAATTTCTTCTAGAGTTCTCTCTTTTACATACATCGAGA	276
Db	291	TCCACTTTTACCTTCATGACGAAATCTTCCAAATGTTCTCCCTTACATCATTTGAAG	350
Qy	277	AAATCTACGGTGCACATTTCTGGTTGGTTCGTCGAACCTTCGGTTAACGGTAGCC	336
Db	351	AAATCTAAGTTTCAATTTTGTTATATAGTTTGGACAAAGCGCTGTGTACCATTTCT	410
Qy	337	GATCTGATTTGATCAGAGAGA--TCTTCTTAAGTGTGATTTTACAGAGAGATGAA	393
Db	411	GAATCCAGCTTTAATTAGGGATTAATTCCTCTGAAATCAATTAATTTAGAGAAATAG	470
Qy	394	GCACACCCCTTGGTTAAACAATTGAAGGCATGACATACTTGTCTCAAAGTGAAAA	453
Db	471	TGCGCGGCGCTGTCAAGAGCTTAGAGTATGTGTTGCTTAGTCTCAAAGGTAAAA	530
Qy	454	TGGGCTCATATCGAAAAATCATTAGCCCTTACTTTCAATATGAGAAATCTTAAATTGCTT	513
Db	531	TGGGCTCACCATAGGAAGATTATTACTCTTCACTTACATTAAGAAATCTCAGGGCTAATG	590
Qy	514	GTACCAAGTTGTATTAAAGTGTGACATATATGGTGAATTAATGTCGATTAAGTATCA	573
Db	591	ATTCCCATGATGGAGAGAGCATGGAAGAGATGTTAACAATATGTGTCAAAAATGTCAAC	650
Qy	574	GAAGACGTGAAGTTAGGTTAGATGTCTATGATGATGTTTTCAGATTTTGACTGAAGATGT	633
Db	651	GCAGCGGAAAAGTAGAGATGGAAGTGTGCAAAATGTTTCCAACTTAGAGAGATGTC	710
Qy	634	ATTATGTAACAACGCTTTTGGAAATGTACATGAAGATGTTGAGACAGTTTTCGACTCA	693
Db	711	ATTACGCGATTTGTGTGGAAACAGCTACGAAAGATGTAAGCCATCTTGAATTTACA	770
Qy	694	GTCCACAATAGCTTTTGTGCTGAACCTTTGAAGAGCTTCAATCTCTGGCTATAGA	753
Db	771	GCACACCAATGTGTTATGCCACCGAAGCTTATCAAAAAGATTCAATCCCGGATACAG	830
Qy	754	TTTTTTCGACAGAGCGAATTTGAAGTCTCGGAATTTAGCAAGAGATAGGAAGTGG	813
Db	831	TTTTTGCTAGTAAAAAGATAGAAATATGTTGGAGATTGATTAACAAGTGCATATCT	890
Qy	814	TTGTGAAGCTGATAGAGCGCGAG 839	
Db	891	TTGATGAAGCTGATTAAGAGAG 916	

RESULT 5	614 bp	MRNA	linear	EST 17-OCT-2002
BUB91384				
LOCUS	PU949G01	Populus petiolaris	Populus tremula	CDNA 5 prime,
DEFINITION	MRNA sequence.			
ACCESSION	BUB91384			
VERSION	BUB91384.1	GI:24102449		
KEYWORDS	EST.			
SOURCE	Populus tremula			
ORGANISM	Populus tremula			
	Eukaryotes; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			
	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;			
	rosids; eucotids I; Malpighiales; Salicaceae; Salicaceae; Populus.			
REFERENCE	1 (bases 1 to 614)			
AUTHORS	Umeberg, P., Bhalerao, R.R., Jansson, S. and Sterky, F.			
TITLE	The poplar tree transcriptome: Analysis of expressed sequence tags			
JOURNAL	from multiple libraries			
COMMENT	Unpublished. (2002)			
	Contact: BHALERAO RUPALI R.			

## FEATURES

upali.bhalerao@plantphys.umu.se.  
Location/Qualifiers

source

1..614  
/organism="Populus tremula"  
/mol\_type="mRNA"  
/db\_xref="taxon:11936"  
/tissue type="petioles"  
/clone\_id="Populus petioles cdna library"

Query Match 19.7% ; Score 307.8 ; DB 13 ; Length 614 ;  
Best Local Similarity 72.4% ; Pred. No. 3e-77 ;  
Matches 399 ; Conservative 0 ; Mismatches 152 ; Indels 0 ; Gaps 0 ;

QY 1013 TCTTGTGATCCAGTCACCGGAGTGCGAGGCCAAACACATGATAGAGGCTCAGAGGCT 1072  
Db 1 TCTTCTAGCAATGATCATCACGGGACGGCTTAAGACATGATAGAGGCTTTGAAGGCT 60  
QY 1073 GCGGCTCACTGATATGCTTACCACAGAACCATGTCTTAAGCTTAAAACGTTGAATGA 1132  
Db 61 GTGATATCACTGAGACGCCACCAAGATATGTTCCAAGGCTTAAGACGCTGACATGA 120  
QY 1133 TCTTGAAGAATCTTTAAGTTGATACCAATAGTAGCTACGATTGACGCGCTAAAT 1192  
Db 121 TCTTAAAGAAATCTTACGCTGTACCCACCAACAATGCCCAATCAGCGGCTCAAG 180  
QY 1193 CGGATGTGAAGCTAGAGGGATACAAATCCCATGTG3ACAGAGCTTCTAATCCATCA 1252  
Db 181 CTGATGTGAGAGCTGAGGGGCTACAAATCCCAAGTGGAGCGAGATCTGATACGATCC 240  
QY 1253 TAGCGGTCCATCAGACCAAGCCATTGGGGGTATAGCTGAGACGATTCAGTCCAGCTC 1312  
Db 241 TGACCTTCATATGATATCAATCAATAGGGGCATATGACAAATGATTCACACCTGCTC 300  
QY 1313 GGTTTGCAGATGAGATGCGCGTGTCTGCCAACAACCCCGTGGCTTCATACGTTTGGCC 1372  
Db 301 GCCTTCGCGATGTGTGGCCGCGCTCCAGCATCAGTGGCGCTTCATTCGTTCCGAGC 360  
QY 1373 TCGGAGTTCGATACGATTTGGTGAAGATCTTCTACTACAGGCCAAATGACATCG 1432  
Db 361 TCGGATCGCAGCTGACATCGGCCAAATTTAGCATCTTTGCAAGCAAATCAAGCTCG 420  
QY 1433 CTGATATGATCCAAAGCTTCACTTCACTGCTCTCTAATTACATGACATGCACTACCG 1452  
Db 421 CTATATGTGTGCAAGATTTCTTTCAAGCTGGCCCCCTGTGATCAACATGACCAAG 480  
QY 1493 TCCTTATGTTGGTTTATCTCAATATGATGTCACCAATCACTTCGCGAGATGACCAATC 1552  
Db 481 TCTCATGTGATCTTCAACCCCAGATAGGCGCACCAATCATCTTCCAAATCTGTCAATC 540  
QY 1553 ATGAGGATGA 1563  
Db 541 ATGAAACATCA 551

RESULT 6  
AY109087 1876 bp mRNA linear HTC 17-OCT-2002  
LOCUS Zea mays PCO144494 mRNA sequence.  
DEFINITION AY109087  
ACCESSION AY109087  
VERSION AY109087.1 GI:21212491  
KEYWORDS HTC.  
SOURCE Zea mays  
ORGANISM Zea mays  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD  
clade; Panicoideae; Andropogoneae; Zea.  
1 (bases 1 to 1876)  
Hainey,C.F., Dolan,M., Miao,G.H., Vogel,J.M., Whiteitt,M.S.,  
Arthur,L.W., Hanafy,M., Morgante,M. and Tingey,S.V.  
Maize Mapping Project/DuPont Consensus Sequences for Design of  
Overgo Probes  
JOURNAL Unpublished (2002)  
REFERENCE 2 (bases 1 to 1876)  
AUTHORS Coe,E.H.



TITLE Direct Submission  
JOURNAL Submitted (25-APR-2002) Maize Mapping Project, University of Missouri, Columbia, MO 65211, USA  
COMMENT If you are interested in getting corresponding physical clones, these are publicly available from ZmDB and may be found by BLAST searching at MSU, maizegap.org; ZmDB, www.zmdb.iastate.edu; TIGR, www.tigr.org; or NCBI, www.ncbi.nlm.nih.gov. When the source of the maize cDNA sequences is either Virginia Walbot, Stanford or Pat Schnable, Iowa State, then clones may be requested from ZmDB: www.zmdb.iastate.edu.

FEATURES  
Source location/Qualifiers  
1..1876  
/organism="Zea mays"  
/mol\_type="mRNA"  
/db\_xref="maizedb:638539"  
/db\_xref="taxon:4577"  
/clone\_lib="Maize Mapping Project/Dupont Cornsensus Library"  
/note="this sequence is part of a project of EST assemblies resulting from the application of public contigs to seed Dupont contigs; this resource was assembled by Dupont as part of a collaboration for the overgo addressing of BACs in conjunction with the Maize Mapping Project"

# ORIGIN

Query Match 18.4%; Score 287.2; DB 11; Length 1876;  
Best Local Similarity 51.4%; Pred. No. 5.1e-71;  
Matches 751; Conservative 0; Mismatches 693; Indels 18; Gaps 3;

QY 94 TGGTGAAGACCAAGAAATGAAACATTTCTTAAACAAGAAATCGAGTCTCTCT 153  
DB TGGTGAAGACCAAGACCAAGAAATGAAACATTTCTTAAACAAGAAATCGAGTCTCTCT 255  
QY 154 TATCATTTCTTCTATGGAATGTTAAAGATCTTGTGAAGATGATGATTAAGTTCTTCT 213  
DB 256 TACCGCTCTCTCAACCGGAGCGTGAAGAAACCTCGCTCAACCGAAGACCGGACG 315  
QY 214 CATCTATGCCC---TTCTCTCAAAATCTCTCTAGAGTTCTCTTTTAACTCAAC 270  
DB 316 AAGCACTGCGCTGGAGATCAGACATCATCCAGCGCTGACCATGTTCCACAC 375  
QY 271 TGGAGAAAAATCTACGCTGCTACATTTGTGTTGTTGCTTCCACTTTCGGTTACG 330  
DB 376 GCGGTGAAGAGAAATGGAGCAAAATTCATTCATTTGCTTGGCCGATTCGAGGGTATA 435  
QY 331 GTAGCGAATCTGATTTGATCAGAGATCTTCTTAAGTGTAGTCTAGAGAAAT 390  
DB 436 ATTCAAGACCCGGAATTAATGAGAGATTTGTCTAAAGTTGGCCACTTTGGCAA 495  
QY 391 GAGCTCAACCTTTGTTAAACACTTGAAGCGGATGACTTAACTGCTCAAAAGTGA 450  
DB 496 CCACGTGTTAGCGGTGGCAAGTACTAGCAATGGCTCGCTAACTGAAGGCGAG 555  
QY 451 AATGGGCTCATCAGAAAAATCATAGCCCTACTTTTCATATGAGAAATCTTAAGTT 510  
DB 556 AATGGGCAAAACAGAGAAATCTCAACCTGCTATCCACATGAGAAAAATTAACGG 615  
QY 511 CTGTACACAGTTGTGTTGAAGTGAATGATATGATGATGATTAATGATGATGAT 570  
DB 616 ATGCTGCTGTTATTTGCTACCTGTTGTCGATATGATTAACAATGAGAAATTCATG 675  
QY 571 TGAAGAAACGTTGAAGTGAATGATGATGATGATGATGATGATGATGATGATGAT 630  
DB 676 TCTTCCAGAAACCTTCTGAATGAGCGCTGGCTGAGTCCAAATCTTACTGAGAT 735  
QY 631 GTTATTAGTAGAAACAGCTTTTGAAGTGAATGATGATGATGATGATGATGATGAT 690  
DB 736 GTTATCTCAAGAACTCTTTTGTGTAACAATGAGAGGAGGAGAAACATCTTCACTA 795  
QY 691 CAGGCTCAAAATGCTTTTGTGCTGAAGCTTTTCAAAAGTCTTCAATCTTCCCTGAT 750  
DB 796 CAGGAGAGAGGCTGAGCCCTTATACATCTTTTCAAAACATTTATATCCAGGCTAC 855

QY 751 AGATTTTCCGACAGAGGAAATTTGAAGTCTCGAAGTTAGACAGAGATTAAGAG 810  
DB 856 TGGCTTTCGCCACCAAAATTAACAAGATGAAGAAATTTGATCGGAGATTCGTTAA 915  
QY 811 TGGTTGTTGAAGTGAATGAGCGCGGAGACAAACGTTATGATGAGAGAGGAGAA 870  
DB 916 ATTCTGATGAAATTAATACAGAAAGAGAGGCTTTTATGATGATGATGATGAT 975  
QY 871 TGTAGAGAGCGCGGAGAGATTTGTTGATGATGATGATGATGATGATGATGATGAT 925  
DB 976 GATGATGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1035  
QY 926 -----TGACGTTTCAAGCATTTGTGAAGAGTGAAGATTTTCTTCCCGGAGAA 978  
DB 1036 CTCGAAATGATACCGAAAGATTAATCGAGAAATGCAAGTTATTTTATCTTGGAGTATG 1095  
QY 979 CAGCAACTTCTAATCTGCTGAGCTGAGACACATCTGCTATCCATGACCCGAGTGG 1038  
DB 1096 GAGCAACATCAGTCTGCTTACATGACGTTAATCTGCTAAGCATGACCTGTAATG 1155  
QY 1039 CAGGCAAGACAGTATGAGTCTCAGGCTGCGGCTCAGTATGATGCTTACCAAG 1098  
DB 1156 CAGAGCAGGCAAGAGAAAGTTTAAACACTTTGG---AATGGGTACACGACCTTT 1212  
QY 1099 GACCATGCTTTAAGCTTAAACGTTGATGATCTTGAACAGATCTTAAAGTTGAT 1158  
DB 1213 GATTAATGAAACCGCTGAAGATTTGTAACCATATTTTATGATGATCTTGAAGTTGAC 1272  
QY 1159 CCACCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1218  
DB 1273 CCACCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1332  
QY 1219 ATCCATGATGAGCAGGAGCTTCAATCCATGATGATGATGATGATGATGATGAT 1278  
DB 1333 TATCTTCAAGCGTGAACCTTCACTGATGATGATGATGATGATGATGATGATGAT 1392  
QY 1279 TGGGTAATGACGTAAGCAATTCATGACGCTGATGATGATGATGATGATGATGAT 1338  
DB 1393 TGGGTAATGACGTAAGCAATTCATGACGCTGATGATGATGATGATGATGATGAT 1452  
QY 1339 GCCAAACCCCGTTGCTTCAATCCGTTTGGCTGAGAGTTGATGATGATGATGATGAT 1398  
DB 1453 ACCAAGATGACGCGGCTTCTTCCATTTGATGATGATGATGATGATGATGATGAT 1512  
QY 1399 AATCTGATATCTTCAAGCAATTCATGACGCTGATGATGATGATGATGATGAT 1458  
DB 1513 AATCTGATATCTTCAAGCAATTCATGACGCTGATGATGATGATGATGATGAT 1572  
QY 1459 CACTGATCTTCAATGACATGACATGACATGACATGACATGACATGACATGACAT 1518  
DB 1573 GAGCTCTATGATCTTCAATGACATGACATGACATGACATGACATGACATGACAT 1632  
QY 1519 GGTGACCAATCAGCTTCCGGA 1540  
DB 1633 GGTGCTCAATCAGTTGANA 1654

RESULT 7  
CG027129/c 698 bp DNA linear GSS 19-AUG-2003  
LOCUS CHG117/c CHG1 cleome hassleriana genomic clone CHG117, genomic  
DEFINITION survey sequence.  
ACCESSION CG027129  
VERSION CG027129.1 GI:33899285  
KEYWORDS GSS.  
SOURCE Cleome hassleriana  
ORGANISM Cleome hassleriana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
rosids; eurosids II; Brassicales; Brassicaceae; Cleome.  
REFERENCE  
Town,C.D., Van Aken,S., Utechtack,T. and Fraser,C.M.



QY 1085 ATGTCCTACCAAGACCATGTCGTTAAGCTTAATAACGTTGATGATCTTTGAACGAGT 1144  
 DB 551 ATCCCTCATTAACACCAAAATTTCCAGCTTAAGACGCTAAGATATATTAATTAAGAAAT 492  
 QY 1145 CTTTAAGGTGATACCAACCAATAGATGATGATGACGCGCTAAATCGATGTGAAG 1204  
 DB 491 CTGTCGGCTATATCCACAGACAGTGGCAGACATACAGACGACCAAGTTGATACAAAT 432  
 QY 1205 TAGAGAGGTACAAATCCCATGTGACAGAGCTTCTATATCCCATATAGCGCTCATC 1264  
 DB 431 TAGGGGATTTACATTAACCTAGAGGAGTGAATCTCTTAATACCAATATGCAATTCATC 372  
 QY 1285 ATGACCAAGCCATTTGGGGTATATACCTGACGAATTCATCCAGCTCGGTTGGGGAGT 1324  
 DB 371 ATGATCAAAACACTATGGGGACAGAGCTTAACGAATTCATCCAGCAAGATTTGGCTAG 312  
 QY 1325 GAGTCCGCGTCTGCCAAACACCCCGTTGCTTATACGCTTTGGCTCGAGATTGCTA 1384  
 DB 311 GAGTGGACACAGCAGACAAAACACCCCATGCGCTTATGCTTTGGCTCGGCGCCGAC 252  
 QY 1385 CATGATGTTGTCAGATCTGCTATCTTCAGGCCAATTTACACTGCTGTATATATTC 1444  
 DB 251 GATGCGTGGCCAAATTTAGCAGTGTACAGCTTAATTAAGCAATGCTATATCTTTC 192  
 QY 1445 AAGGCTTACCTTACCTTGGCTCTCTATCTATACAGATGACCTACGCTCTTATGTTGC 1504  
 DB 191 AAGGCTTTTCGTTTATCTTTCTCCATTAACAGCATGCTCCCATATTTGATGCTCT 132  
 QY 1505 TTTATCTTCAACATGTTGACCAATCACCTTCCGAGATTG 1545  
 DB 131 TATGTCCAAATATGCTGCTCTATTAATTCACAAAGTTG 91

RESULT 9  
 BH597310 817 bp DNA linear GSS 15-DEC-2001  
 LOCUS BH597310  
 DEFINITION BH597310 BOLT Brassica oleracea genomic clone BH597310, genomic survey sequence.  
 ACCESSION BH597310 GI:17849762  
 VERSION BH597310  
 KEYWORDS GSS.  
 SOURCE Brassica oleracea  
 ORGANISM Brassica oleracea  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eustosids II; Brassicales; Brassicaceae; Brassica.  
 REFERENCE 1 (bases 1 to 817)  
 AUTHORS Town, C.D., Van Aken, S., Uteirback, T., Koo, H. and Fraser, C.M.  
 TITLE Whole genome shotgun sequencing of Brassica oleracea  
 JOURNAL Unpublished (2001)  
 COMMENT Other GSSs: BH597310  
 CONTACT: Chris Town  
 TIGR  
 9712 Medical Center Drive, Rockville, MD 20850, USA.  
 Tel: 301-838-3523  
 Fax: 301-838-0208  
 Email: cdtown@tigr.org  
 DNA is from a doubled haploid provided by Tom Osborn.  
 Seq primer: TR  
 Class: sheared ends.

FEATURES  
 source  
 1..817  
 Location/Qualifiers  
 /organism="Brassica oleracea"  
 /mol\_type="genomic DNA"  
 /strain="TO1000DH3"  
 /db\_xref="taxon:3712"  
 /clone="BH597310"  
 /clone\_1="BH597310"  
 /note="Vector: PHOS1; Site 1: BstXI; 2-3 kb sheared genomic DNA inserted into PHOS1 using BstXI linkers"

ORIGIN

Query Match 16.6%; Score 259.6; DB 28; Length 817;  
 Best Local Similarity 83.9%; Pred. No. 3e-63;  
 Matches 317; Conservative 0; Mismatches 59; Indels 2; Gaps 2;

QY 1122 GTTGAATGATCTTTGAACAGATCTTTAAGTTGATATCCCAATATGATGATGATG 1181  
 DB 441 GTTGGGATGATCTTTGAACAGATCTTTAAGTTGATATCCCAATATGATGATGATG 500  
 QY 1182 ACGGCTTAATCGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1241  
 DB 501 ACGGCTTAATCGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 560  
 QY 1242 AATCCCATCATATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1301  
 DB 561 GATCCCATCATATGATGATGATGATGATGATGATGATGATGATGATGATGATG 620  
 QY 1302 CATTCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1361  
 DB 621 CATTCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 680  
 QY 1362 ACCGTTGGCTCGAGTGTGATGATGATGATGATGATGATGATGATGATGATG 1421  
 DB 681 ACCGTTGGCTCGAGTGTGATGATGATGATGATGATGATGATGATGATGATG 739  
 QY 1422 ATGACACTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1480  
 DB 740 ACTGACCTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 799  
 QY 1481 ATGACACTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1540  
 DB 800 ACGACCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 817

RESULT 10  
 BQ297122 493 bp mRNA linear EST 16-MAY-2002  
 LOCUS BQ297122  
 DEFINITION BQ297122 Y1 Gm-c1054 Glycine max cDNA clone SOYBEAN CLONE ID: Gm-c1054-8364 5' similar to TR:048786 048786 PUTATIVE CYTOCHROME P450. ;, mRNA sequence.  
 ACCESSION BQ297122 GI:20812644  
 VERSION BQ297122  
 KEYWORDS EST.  
 SOURCE Glycine max (soybean)  
 ORGANISM Glycine max (soybean)  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eustosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.

REFERENCE 1 (bases 1 to 493)  
 AUTHORS Shoemaker, R., Keim, P., Vodkin, L., Epeiding, J., Coryell, V., Khanna, A., Bolla, B., Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T., Underwood, K., Stepien, M., Theising, B., Allen, M., Bowers, J., Peterson, B., Swaller, T., Gibbons, M., Page, D., Harvey, N., Schurk, R., Riltter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R., Waterston, R. and Wilson, R.  
 TITLE Public Soybean EST Project  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Shoemaker R/Public Soybean EST Project  
 Public Soybean EST Project  
 Washington University School of Medicine  
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA  
 Tel: 314 286 1800  
 Fax: 314 286 1810  
 Email: est@wustl.edu  
 This clone is available through: Reggen, Invitrogen Corp. 2130  
 South Memorial Parkway Huntville, AL 35801 For further information  
 call: (800)-533-4363 or contact: cst@reggen.com web site:  
 www.reggen.com  
 Seq primer: -40RP from Gibco  
 High quality sequence stop: 425.  
 Location/Qualifiers  
 1..493  
 /organism="Glycine max"

FEATURES  
 source  
 1..493  
 Location/Qualifiers  
 /organism="Glycine max"

/mol\_type="mRNA"  
/db\_xref="taxon:3947"  
/clone="SOBESAN CLONE ID: Gm-cl054-6364"  
/issue\_type="leaf", 3 week old, greenhouse grown"  
/lab\_host="DH10B"  
/clone\_lib="Gm-cl054"  
/note="Vector: pBluescript II SK+; Site 1: EcoRI; Site 2: XhoI; The Harosoy NIL was constructed and seed was provided by Dr. J. Specht, University of Nebraska (Shoemaker and Specht, 1995). The cDNA library was constructed from mRNA isolated from leaf tissue at various developmental stages of 3 week old greenhouse grown plants. Complementary DNA was synthesized from mRNA using a primer consisting of a poly(AT) sequence with a XhoI restriction site and a 3' anchor. EcoRI adapters were ligated to the blunt-ended cDNA fragments followed by XhoI digestion. The cDNA fragments were directionally cloned into the EcoRI-XhoI restriction site of the pBluescript vector. The ligated cDNA fragments were transformed into DH10B host cells (GibcoBRL). This library was constructed in cooperation with Dr. Paul Keim's laboratory at Northern Arizona University."

## ORIGIN

Query Match 16.4%; Score 256; DB 13; Length 493;  
Best Local Similarity 70.3%; Pred. No. 2.5e-62;  
Matches 343; Conservative 0; Mismatches 145; Indels 0; Gaps 0;

QY 1057 GAGTCTCAGGCTCGCGCTCAGCGTGAATGCCACAGAGACATGCTTAACTT 1116  
Db 2 GAGTCTCAGGCTCGCGCTCAGCGTGAATGCCACAGAGACATGCTTAACTT 61  
QY 1117 AAAACCTTGAGTATGATCTTGAACGAGTCTTAACTTGAATCCCAATAGTACTAG 1176  
Db 62 AGGACGCTGAGCATGATTTGAAAGATCAGTCACTGACCTCCGACGATCCGCGCA 121  
QY 1177 ATTGACGCGCTAATATGATGTAAGAGGCTAGCAAAATCCCATGTCGCGAG 1236  
Db 122 ATCAGAGGCGCAAGGAGGAGTGAATTTGGGGGATACAAATACCATGTCGCGAG 181  
QY 1237 CTTCTATCCCAATCATAGCGGCTCATCATGACCAACCATTTGGGTAATGACGTGA 1296  
Db 182 CTGCTATTCATTCCTGCGCGCTTCAATCAGATCAAGCAATATGGGGAGACATGTGA 241  
QY 1297 GAATTCATTCAGTCTGCTTGGGAGTGAAGTCCCGGCTGCCAAACCCCGTTGGC 1356  
Db 242 GAGTTCATTCGCGGCGCTTTCAGCGATGATGTCGCGCGCGGGAGATCATTTGGCG 301  
QY 1357 TTCTATCGCTTGGCTCGGAGTTCGATCATGATGTCAGATCTTCTATTTACTTTCG 1416  
Db 302 TTCTATCGCTTGGCTCGGAGTTCGATCATGATGTCAGATCTTCTATTTACTTTCG 361  
QY 1417 GCCAATTCAGATTCGCTGTAATGATTCACAGCTTTCATCTTACTGCTCTACTTAT 1476  
Db 362 ACAAACCTGCGCTCGGAGATTCATCTGACAGCTTTCATCTTACTGCTCTACTTAT 421  
QY 1477 CAGATGACCTACGCTCTATATGTTCTTATCTTCAATCATGTCAGATTCACCTTTC 1536  
Db 422 CAGATGACCTACGCTCTATATGTTCTTATCTTCAATCATGTCAGATTCACCTTTC 481  
QY 1537 CGGAGATT 1544  
Db 482 CAACAT 489

RESULT 11 303 bp DNA linear GSS 21-NOV-1997  
B62694/c  
LOCUS T17P24TF.1 TAMU Arabidopsis thaliana genomic clone T17P24, genomic  
DEFINITION survey sequence.  
ACCESSION B62694  
VERSION B62694.1 GI:2629456  
KEYWORDS GSS.

SOURCE Arabidopsis thaliana (thale cress)  
ORGANISM Arabidopsis thaliana  
REFERENCE 1 (bases 1 to 303)  
AUTHORS Rounsley, S.D., Field, C.E., Bass, S., Linher, K., Linher, K., Golden, K., Berry, K., Granger, D., Sun, E., Wilde, C., Adams, M.D. and Venter, J.C.  
TITLE A BAC End Sequence Database for Identifying Minimal Overlaps in Arabidopsis Genomic Sequencing. Update 3  
JOURNAL Unpublished (1997)  
COMMENT Other\_GSS: T17P24TF T17P24TRB  
Contact: Steve Rounsley  
Department of Eukaryotic Genomics  
The Institute for Genomic Research  
9712 Medical Center Dr., Rockville, MD 20850, USA  
Tel: 301 838 0200  
Fax: 301 838 0208  
Email: rounsley@igr.org  
Seq primer: M13-21  
Class: BAC ends  
High quality sequence stop: 303.  
FEATURES  
source  
1..303  
/organism="Arabidopsis thaliana"  
/mol\_type="genomic DNA"  
/strain="Columbia"  
/db\_xref="taxon:3702"  
/clone="T17P24"  
/sex="hermaphrodite"  
/clone\_lib="TAMU"  
/note="Vector: pBluescript II SK+; Site 1: HindIII; Site 2: HindIII; Produced by Rod Wing"

## ORIGIN

Query Match 15.6%; Score 244; DB 28; Length 303;  
Best Local Similarity 98.0%; Pred. No. 5.5e-59;  
Matches 247; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1312 CGGTTTGGCGATGAGTGGCGGCTGCTGCGCAACGCCGTTGGCTATACCGTTGGC 1371  
Db 303 CGTTTCCGATGAGTGGCGGCTGCTGCGCAACGCCGTTGGCTATACCGTTGGC 244  
QY 1372 CTGCGAGTTCGATACGATTCGATGATGATTCGATTCGATTCGATTCGATTCGATTC 1431  
Db 243 CTGCGAGTTCGATACGATTCGATGATGATTCGATTCGATTCGATTCGATTCGATTC 184  
QY 1432 GCTGTATGATTCACAGCTTTCATCTTACTGCTCTACTATACAGATGACCTTAC 1491  
Db 183 GCTGTATGATTCACAGCTTTCATCTTACTGCTCTACTATACAGATGACCTTAC 124  
QY 1492 GTCTTATGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTC 1551  
Db 123 GTCTTATGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTC 64  
QY 1552 CATGAGATTGA 1563  
Db 63 CATGAGATTGA 52

RESULT 12 834 bp DNA linear GSS 20-AUG-2003  
CG112869/c  
LOCUS PUIAK51TB ZM\_0.6\_1.0\_KB Zea mays genomic clone ZM06T053805,  
DEFINITION genomic survey sequence.  
ACCESSION CG112869  
VERSION CG112869.1 GI:33996306  
KEYWORDS GSS.  
SOURCE Zea mays  
ORGANISM Zea mays  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD

Clade: Panicoidae; Andropogoneae; Zea.  
1 (bases 1 to 834)  
Whitehead, C.A., Quackenbush, J., Van Aken, S., Utterback, T.,  
Resnick, A., Fraser, C.M., Yuan, Y., San Miguel, P., Ma, J. and  
Bennetzen, J.

TITLE  
JOURNAL  
COMMENT  
Maize Genomics Consortium  
Unpublished (2003)  
Other GSSs: PUIK51TD  
Contact: Cathy Whitehead  
TIGR

9712 Medical Center Drive, Rockville, MD 20850, USA  
Tel: 301-838-5843  
Fax: 301-838-0208  
Email: whitehead@tigr.org  
Seq primer: TR

FEATURES  
source  
Location/Qualifiers  
1..834  
/organism="Zea mays"  
/mol\_type="genomic DNA"  
/strain="B73"  
/db\_xref="taxon:4577"  
/clone="ZMMB180538J05"  
/clone\_id="ZM\_0.6\_1.0\_KB"  
/note="Vector: pCR4-TOPO; Site\_1: EcoRI; 0.6-1.0 kb high  
cot selected genomic DNA library"

## ORIGIN

Query Match 15.5%; Score 242.4; DB 29; Length 834;  
Best Local Similarity 61.9%; Pred. No. 2.9e-56; Indels 0; Gaps 0;  
Matches 384; Conservative 0; Mismatches 236; Indels 0; Gaps 0;

926 TGACGGTTCAGGACCTTGTGAGAGAGGTAAAGCTTTTCTTTCGCCGGAGAAACAGACAA 985  
631 TGCGCGCGAGACAGATCATCGAGAGATCAAGAACTTCTTTCGCCGGAGAGAGAGCG 572  
986 CTTTAAATCTGCTGACGTGAGACACCATCTTGTCTATCCATGACCCGGAGTGGCAGGCCA 1045  
571 TCACCAAGTTTGTCTCCTCGGCGCACCGCTGCGCTTACCATGACACAGAAATGGACAGCC 512  
1046 AAGCAGTGTAGAGTCTCTCAGGGTCTCGGCTCAGCGATGCTCCCTACCAAGACCATG 1105  
511 GCGCTCGCAGAGATCATGAGTCTCGGCGACCGGCGCGTCCCAACCAAGACACACA 452  
1106 TCGTTAAGCTTAAACGTTGAGTATGATCTTGAACGAGTCTTTAAGTTGTATCCACAA 1165  
451 TACCAAGGCTCAAGCGCTCGGAGTGTATCAACAGACCGCTCAGGCTTACCGCGCG 392  
1166 TGTAGCTACGATTTGAGCGCGCTAATCGAATGTAAGCTAGAGGGTACAAATCCCAT 1225  
391 CGGTGCCATGATCCGAAAGCGGAGCGGAGCTGAGACTCGGCGCTGCTGTGCGG 332  
1226 GTGGACCGAGCTTAAATCCCATCATAGCGGTCCATCATGACCAAGCATTTGGGCTA 1285  
331 CGGGACCGAGATCATGATCCCATCATGCGGTACACACAGACCGCGCTGTGGGCG 272  
1286 ATGAGCTGACGAATTCATCCAGCTCGGTTTGGGAGTGGAGTCCGCGGTGCTGCCAAC 1345  
271 ACGAGCGCACGAGATTCACCCCGCGGCTTTCGCGAGAGATGAGCCGCGGCGGCG 212  
1346 ACCCGGTGCTTCAATACCGTTTGGCTTCGAGTTTCATGATGATGATGATGATGATG 1405  
211 AGAGATGCGCTTCTTCCGTTCCGCGGCGGCGCGCGGTGTGATCGGCGAGTACCTGG 152  
1406 CTAATCTTACGCGCAATTTGACATCGCTGTAATGATCAACGCTTCACTTCACTTGG 1465  
151 CGCTCATGAGCGCAAGATCGCTTGGCTACGATGCTCAAGCGCTGCAAGTTCGCGCTGT 92  
1466 CTCCTACTTATGACATGACATGACATGACATGACATGACATGACATGACATGACATG 1525  
91 CGCGCGCTGATGAGCGCGCGCGGAGGATGATGATGATGATGATGATGATGATGATGATG 32

Db 31 CGTCACTTCGCCGACATG 12

RESULT 13  
AI104671  
LOCUS  
DEFINITION  
Zea mays PC0144495 mRNA sequence.  
AY104671  
VERSION  
AY104671.1 GI:21207749  
KEYWORDS  
SOURCE  
ORGANISM  
Zea mays  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD  
clade: Panicoidae; Andropogoneae; Zea.  
1 (bases 1 to 1567)  
Haine, C.F., Doan, M., Mao, G.H., Vogel, J.M., Whitesitt, M.S.,  
Arthur, L.W., Hanaley, W., Morgante, M. and Tingey, S.V. Design of  
Maize Mapping Project/Dupont Consensus Sequences for Design of  
Overgo Probes  
Unpublished (2002)  
2 (bases 1 to 1567)  
Coe, E.H.  
Direct Submission  
Submitted (25-APR-2002) Maize Mapping Project, University of  
Missouri, Columbia, MO 65211, USA

COMMENT  
If you are interested in getting corresponding physical clones,  
these are publicly available from ZMDB and may be found by BLAST  
searching at MSU, maizegap.org; ZMDB, www.zmdb.iastate.edu; TIGR,  
www.tigr.org; or NCBI, www.ncbi.nlm.nih.gov. When the source of the  
maize cDNA sequences is either Virginia Walbot, Stanford or Pat  
Schubert, Iowa State, then clones may be requested from ZMDB:  
www.zmdb.iastate.edu.  
Location/Qualifiers  
1..1567  
/organism="Zea mays"  
/mol\_type="mRNA"  
/db\_xref="maizedb:638540"  
/db\_xref="taxon:4577"  
/clone\_id="Maize Mapping Project/Dupont Consensus  
Library"  
/note="this sequence is part of a project of EST  
assemblies resulting from the application of public  
contigs to seed Dupont contigs; this resource was  
assembled by Dupont as part of a collaboration for the  
overgo addressing of BACs in conjunction with the Maize  
Mapping Project"

## FEATURES

source

## ORIGIN

Query Match 15.0%; Score 234; DB 11; Length 1567;  
Best Local Similarity 50.2%; Pred. No. 1.1e-55;  
Matches 632; Conservative 0; Mismatches 620; Indels 6; Gaps 2;

284 ACGGTGCTACATTTCTGTTTGTGTTGGTCCACTTCCGTTAAACGTTAGCGGATCGT 343  
101 ACGGAGAGACAGAGATGCTCGGTTGCGGCCACGCCCAAGAGTACATGTCGACCCGG 160  
344 ATTGATCAGAGAGATCTTCTTAAGTCTGAGTTCTACGAGAGATGAGAGTCAACCTT 403  
161 ACGTGCACAGAGAGTGTCTTCAAGTTCGACATGCTGCGGCTTCAAGAGTCAAGTCCGG 220  
404 TGGTTAAACAACTTGAAGCGCATGACTTAACTTCAAGTTCGAAGTTCGAAGTTCGATC 463  
221 CGCTGTCAAGATGTGGGCTCGCGCGTGGCGCACACAGAGGAGAGAGTGGTCAAGC 280  
464 ATCGAAATATGATGACCTTCTTATGAGAGATCTTAAGTCTTGTACAGTTG 523  
281 ACCGAGAGATCTCAACCTTGTGCTTCACTTGAAGAGTCAACGAGATGTCGCGCGT 340  
524 TGTGAAGAGTGTGATGATATGATGATGATGATGATGATGATGATGATGATGATGATG 580  
341 TTCTACGCTGTCAGAGAGCTGTGTGACAGTGTGGGCGGCGAGATCTTGTGCTCGAGC 400

Mon May 24 08:10:09 2004

us-09-992-901-1.rst

Page 10

QY	581	GTAAATTTAGGATAGTATGCTATAGATGGTTTCAGATTTGACTGAAGATGTTATTA	640
Dp	401	GCTGTCGAGCTGGAGCGTCTGGCCGAGCTGCAGAACTGACGGGAGCATTTTCTC	460
QY	641	GAAACACTTTTGGAACTAGCTATGAAGTGGTCAGACAGTTTTTGCATTTCAAGCTCAAC	700
Dp	461	GCACCGCTGTACAGCACAGACTACACGGAAGGAAAGAACTTCCAGTCAAGSCGAGC	520
QY	701	AAATGCTTTTGTGTGTAAGCTTTTCAAAAGCTTCATTCCTGGCTATAGATTTTTC	760
Dp	521	AACTTCGCTGTGATGACCAACATTCGGAAGATCATGATCCCGACTACATGTACTTGC	580
QY	761	CGACAAAGAGGAATTTGAAGTCTCGGAATTAGCACAAGATTAAGAAAGTGGTTGTA	820
Dp	581	CTACCGGAAACACAGAAATGATGCCAGAAACACAGGAGGT--GGATCAATCTTGC	637
QY	821	AGCTGATGAGCCGCGAGACAAACCGTATGATGAGAAAGGGAAGATGAAGAGC	880
Dp	638	GAGAGTAATTCGGAAAGAGATCCAGCGATGAAACAGACGAGCGCACAAAGCAGCA	697
QY	881	CGCGCGCAAGGATTTGGATTTAATGATTCAGCGAAGAAATGACGTTTCAGACA	940
Dp	698	TGCTGGGTCTACGTCTGAGACAAACATGAAACGCACTGGGAGATGACATCGAAGC	757
QY	941	TTGTGAGAGATGTAAAAGCTTTTCTTCGCGGGAACACAGCACTTCTAATCTGTGA	1000
Dp	758	TGATTCAGAGAGTGCAGAGGTCTTACTTCATCCGAGATGAGAACACAGTGGTGTCTCA	817
QY	1001	CGAGACGACCAATCTTGCTATCAATGACACCGGAGTGGAGCGCAAGACATGATAGAG	1060
Dp	818	CGTGAACGATGTGTGTGCTCAGACATGCACCCGAGTGGCAGAACCCGCGAGAGAGG	877
QY	1061	TCCTCAGGGTCTGCGGCTCACGTATGTCCCTACAAAGACATGTCTTAAGCTTAAA	1120
Dp	878	TGACGGCTCTGTTCGACAGGAGCAGCAAGCCCGAGTATGACGGCTTACCGCTCAAG	937
QY	1121	CGTTGAGTATGATCTTTGAACAGCTTTTAAAGTGTATCCACCAATAGTAGTAGATTC	1180
Dp	938	TGTGTACATGGTGCTCTACAGAGTGCTCCGACTGTACCGCGGCAACCTCCGTCTCC	997
QY	1181	GAGCGCTTAATCGATGTGAAGCTTAGAGGGGTCAAAATCCCATGTGGACGAGACTC	1240
Dp	998	GACAAACGTACAGAGATGAGAGGTGGAGGCGTCACTAACCCCGCGGATGATCTGG	1057
QY	1241	TAACTCCAAATCAAGCGGTCCATCAATGACCAAGCATTTGGGGTAAATGACGTAAAGAT	1300
Dp	1058	AGCTGCCGCTGCTCATCCACCAACCGGACATCTGGGAGGCGACGCGCCGAGT	1117
QY	1301	TCAATCCACTGGTGTGGATGGATGGAAGTGGCGGCTGTGCACAAACCCCGTGGCTCA	1360
Dp	1118	TTAGGCCCATAGGTTCTCCAGAGGCGTGTCCAGGCGCTCMAAGACCCCGGCGTTC	1177
QY	1361	TACGTTTGGCTTCGAGGTTGTATACATGATGGTCAAACTTGTCTAATCTTCAAGCCA	1420
Dp	1178	TCTGTGTGGTTGAGCCGCGGCACTGTGCATCGGACAGAACTTCGCGTGTGGAGGCA	1237
QY	1421	AATTAGACATCGGTGAATGATCCAAAGCTTCACCTTCACTTGGGCTCTACTTTACG	1480
Dp	1238	AGATGGCGCTTCGATGATCTCTCCAGCGGTTTAGGTTGGGCGTCCGCGCATCTATGCT	1297
QY	1481	ATGCACCTACCGCTTATGTGCTTATATCTCAACATGGTGAACCAATCACTTCG	1538
Dp	1298	ATGCGCGCATACCATGATTAACCTTATGACGAGTGTGCGAGCTTAAGCTCAG	1355

RESULT 14	LOCUS	DEFINITION	ACCESSION
AV417119	AV417119	424 bp mRNA	linear
	AV417119	Lotus japonicus young plants (two-week old) Lotus	EST 28-JUL-2003
		coronilicatus var. japonicus cDNA clone M94135C02_r5, mRNA	
		sequence.	
	AV417119		

VERSION	AV417119.1	GI:7746297
KEYWORDS	EST.	
SOURCE	Lotus corniculatus var. japonicus (Lotus japonicus)	
ORGANISM	Lotus corniculatus var. japonicus	
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosids; eustoside I; Fabales; Fabaceae; Papilionoideae; Lotaeae; Lotus.	
REFERENCE	1 (bases 1 to 424)	
AUTHORS	Asanatsu, E., Nakamura, Y., Sato, S. and Tabata, S.	
TITLE	Generation of 7137 non-redundant expressed sequence tags from a legume, Lotus japonicus	
JOURNAL	DNA Res. 7 (2), 127-130 (2000)	
MEDLINE	20277479	
PUBMED	10819328	
COMMENT	Contact: Yasukazu Nakamura The First Laboratory for Plant Gene Research Kazusa DNA Research Institute Yana 1532-3, Kisarazu, Chiba 297-0812, Japan Email: ynakamu@kazusa.or.jp, URL: <a href="http://www.kazusa.or.jp/en/plant/">http://www.kazusa.or.jp/en/plant/</a> location/Qualifiers	
FEATURES	1..424	
source		

```

/organism="Lotus corniculatus var. japonicus"
/mol_type="mRNA"
/isolate="Miyakojima MG-20"
/db_xref="taxon:34305"
/clone="MM139C03_xr"
/dev_stage="young plants (two-week old)"
/clone_1b="lotus japonicus young plants (two-week old)"
/notes="vector: pBluescriptII SK-, Site_1: EcoRI, Site_2:
XhoI, synonym: Lotus japonicus"

```

Query Match	Similarity	Score	DB	Length
Best Local	14.9%	232.8	DB 9	424
Matches	300	Conservative	0	Mismatches 112; Indels 0; Gaps 0

  

Query	434	TTAGTCTCAAAAGTCAAAAATGCGCTCATCATGAAAATCAATTAAGCCTACTTTTCATA	493
Db	1	TTAGCCTCAAAAGGAGAAAATAGGGCTCACCAAGAAAATTAATATCACCCACTTTTACA	60
Qy	494	TGAGAGATCTTAAGTGGCTTGTACACAGTGTGTGAAGGTGACTGATAGTGTGATA	553
Db	61	TGAAAATCTTCAAATTTGCTGTGTGCCGTATGCGACACACGCTGTGGAGATGTTGGAGA	120
Qy	554	AAATGTCGAGTAAGTTATCAGAAAAACGGTAAGTTGAGTGAATGTCATATGAGTGCTTC	613
Db	121	AGTGGGCTGCATATGGGTGGCAAAAGGATGAGTAATAGAAATTTCTGTAGTGCTTTC	180
Qy	614	AAATTTGACGGAAGATGTTTATGTTGAAACACTTTTGGAGTAGCTATGGAAGATGTC	673
Db	181	AAACTTAAACGAAATGTTATCTATAGACCCGGTTTGGACACGCTACGAAATGCGCA	240
Qy	674	GAGCAGTTTTCGACTTCACAGCTCAACAAATGCTCTTTGTGCTGAAGCTTTTCAAAAAG	733
Db	241	AAGCATTTTCCGCTTCACGCCAACCCAAAGAAATGATCTTAGTGCTATGTTTTCAAAAG	300
Qy	734	TCTTCAATTCGCGCTATAGATTTTTTTTCCGACAAAGAGGAATTTGAAGCTTCGGAAGTGA	793
Db	301	TTTTTATTCGAGTTACAGATTTCTTCCCAAGAGGAGAAATATAAGTTTGGAAAATCGG	360
Qy	794	ACAAGAGATGAAGAGTCGTTGTAAGCTGATAGAGCGGCGGAGACAAAA	845
Db	361	ACAAGGAATTAAGAAATCACTAGTAGAGCTTTTGAAGAAAGAGAGAGAA	412

RESULT 15	CC336779/c	815 bp	DNA	linear	GSS 16-MAY-2003
LOCUS	CC336779				
DEFINITION	CGUHE22TIV.ZM.0.7.1.5_KB				
ACCESION	CC336779				
VERSION	CC336779.1				
					genomic survey sequence.
					clone ZMNMEN0392E01,
					GI:30806192

Mon May 24 08:10:09 2004

us-09-992-901-1.rst

Page 11

KEYWORDS  
SOURCE

Zea mays  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD  
Clade; Panicoidae; Andropogoneae; Zea.  
1 (bases 1 to 815)  
White, C.A., Quackenbush, J., Van Aken, S., Utecher, T.,  
Rendick, A., Fraser, C.M., Budiman, M.A., Bedell, J.A., Rohlfing, T.,  
Clerk, R.W., Numberg, A., Robbins, D. and Lakey, N.  
Consortium for Maize Genomics  
Unpublished (2002)  
Contact: Cathy White, cwhite@maizegenomics.org

TITLE  
JOURNAL  
COMMENT

9712 Medical Center Drive, Rockville, MD 20850, USA  
Tel: 301-838-5843  
Fax: 301-838-0208  
Email: white@maizegenomics.org  
Seq primer: TF  
Class: sheared ends.

FEATURES  
SOURCE

1..815  
/organism="Zea mays"  
/mol\_type="genomic DNA"  
/strain="B73"  
/db\_xref="taxon:4577"  
/clone="ZMMEM0392E01"  
/note="Vector: pBSK-; Site 1: HincII; 0.7-1.5 kb  
methylation filtered genomic DNA library"

ORIGIN

Query Match 14.0%; Score 219.6; DB 28; Length 815;  
Best Local Similarity 57.3%; Pred. No. 1.1e-51;  
Matches 442; Conservative 0; Mismatches 319; Indels 11; Gaps 2;

727 CAAAAGCTTCATTCCTCGGTATAGATTTTCCGACAGAGGAAATTGAAGTCTCG 785  
773 CAGCAAGCTTCCTCCCGGCTACAGGTTCTCTGCCACAGAGAAACAGCGCTGTG 714  
786 GAAGTTAGACAGAGATAGAGAGTCTGTTGAGCTGATAGAGCGCGAGACAAAA 845  
713 GAGCTGACAGAGAGATCAGAGAGCTGTGCTGCAAGCTCGTCCGAGCTGAGAGCG 654  
846 CGCTATAGATGAGAGAGAGAGAGATGTAAGAGCGCGCGGAGAGATTGTTGGATT 905  
653 CGGACGCGGTATCCAGACACAGAGCGCGGAGACAGGCGCGCGCGGAGATGA 594  
906 AATGATTCAG 955  
593 GGGATTCATGAGCTTCATGAGCGCGCGGAGAGAGAGAGAGAGAGAGAGATCCA 534  
956 AAGCTTTTCTTCGCGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1015  
533 AGAATCTTCTTCGCGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 474  
1016 TGCTATCATGACCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1075  
473 CGCTAGCCATGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 414  
1076 GCTCAGTATGCTCCCTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1135  
413 GAGCGCGGCGGCTCCACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 354  
1136 TGAAGAGCTTTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1195  
353 TCAACGAG 294  
1196 ATGTGAAGCTAG 1255  
293 AGGTGAAGCTAG 234

233 CGGTACACACAG 174  
1316 TTGCGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1375  
173 TCGGAG 114  
1376 GAGTGTACATGATGATGATGATGATGATGATGATGATGATGATGATG 1435  
113 GCGGCGGAG 54  
1436 TAATGATCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1487  
53 CGGTGTGAG 2

Search completed: May 22, 2004, 20:13:11  
Job time : 4347 secs

Mon May 24 08:10:08 2004

us-09-992-901-1.rn1

Page 1

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 22, 2004, 16:23:56 ; Search time 133 Seconds  
(without alignments)  
6521.716 Million cell updates/sec

Title: US-09-992-901-1  
Perfect score: 1563  
Sequence: 1 atggaagaagaagtagcag.....tgaccatcatgagatcga 1563

Scoring table: IDENTITY NUC  
Gapop 10.0, Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :  
1: /cgn2\_6/ptodata/2/ina/5A\_COMB.seq:\*  
2: /cgn2\_6/ptodata/2/ina/6A\_COMB.seq:\*  
3: /cgn2\_6/ptodata/2/ina/6B\_COMB.seq:\*  
4: /cgn2\_6/ptodata/2/ina/6B\_COMB.seq:\*  
5: /cgn2\_6/ptodata/2/ina/PTCUS\_COMB.seq:\*  
6: /cgn2\_6/ptodata/2/ina/backfile1.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1563	100.0	1563	4	US-09-527-073-1 Sequence 1, Appl1
2	270.4	17.3	1584	4	US-09-351-229-1 Sequence 1, Appl1
3	184	11.8	588	4	US-09-527-073-16 Sequence 16, Appl1
4	75.8	4.8	612	4	US-09-615-192A-101 Sequence 101, Appl1
5	75.8	4.8	612	4	US-09-169-789-101 Sequence 101, Appl1
6	73.4	4.7	398	4	US-09-615-192A-169 Sequence 169, Appl1
7	73.4	4.7	398	4	US-09-169-789-169 Sequence 169, Appl1
8	68.6	4.4	285	4	US-09-313-294A-2775 Sequence 2775, Appl1
9	53	3.4	7218	1	US-08-232-463-14 Sequence 14, Appl1
10	48	3.1	933	4	US-09-118-554-55 Sequence 55, Appl1
11	48	3.1	933	4	US-09-118-627-55 Sequence 55, Appl1
12	48	3.1	933	4	US-09-602-877A-55 Sequence 55, Appl1
13	48	3.1	940	4	US-09-602-877A-102 Sequence 102, Appl1
14	46.2	3.0	1708	3	US-08-991-677-1 Sequence 1, Appl1
15	45.8	2.9	4403765	3	US-09-103-840A-2 Sequence 2, Appl1
16	45.8	2.9	4411529	3	US-09-103-840A-2 Sequence 1, Appl1
17	44	2.8	2382	4	US-09-023-655-1436 Sequence 1436, Appl1
18	43.6	2.8	1512	4	US-08-277-031B-4 Sequence 4, Appl1
19	43.6	2.8	2059	4	US-09-023-655-1062 Sequence 1062, Appl1
20	43.6	2.8	2759	4	US-09-144-367-1 Sequence 1, Appl1
21	43	2.8	1515	4	US-09-583-447A-3 Sequence 3, Appl1
22	41.6	2.7	1725	4	US-08-882-164D-31 Sequence 31, Appl1
23	41.6	2.7	4411529	3	US-09-103-840A-1 Sequence 1, Appl1
24	41.4	2.6	1893	1	US-08-533-065B-1 Sequence 1, Appl1
25	41.4	2.6	2327	4	US-09-852-067-1 Sequence 1, Appl1
26	40.2	2.6	1634	4	US-09-126-420A-2 Sequence 2, Appl1
27	40.2	2.6	1737	4	US-09-126-420A-1 Sequence 1, Appl1

28	40	2.6	4403765	3	US-09-103-840A-2 Sequence 2, Appl1
29	38.4	2.5	1932	4	US-09-518-386B-2 Sequence 2, Appl1
30	37.8	2.4	936	2	US-08-557-309B-9 Sequence 9, Appl1
31	37.8	2.4	936	3	US-08-834-306-9 Sequence 9, Appl1
32	37.8	2.4	936	3	US-08-993-674A-9 Sequence 9, Appl1
33	37.8	2.4	936	4	US-08-256-976-9 Sequence 9, Appl1
34	37.6	2.4	289	3	US-08-007-005-17 Sequence 17, Appl1
35	37.6	2.4	289	3	US-09-244-796-17 Sequence 17, Appl1
36	37.6	2.4	7218	1	US-08-232-463-14 Sequence 14, Appl1
37	37.4	2.4	1665	3	US-08-881-784-8 Sequence 8, Appl1
38	37.4	2.4	1665	3	US-09-292-768-3 Sequence 3, Appl1
39	37.4	2.4	1665	3	US-09-292-768-67 Sequence 67, Appl1
40	37.4	2.4	1665	3	US-08-292-768-67 Sequence 67, Appl1
41	37.4	2.4	3965	4	US-08-518-386B-4 Sequence 4, Appl1
42	37.2	2.4	274	4	US-09-313-294A-1017 Sequence 1017, Appl1
43	37.2	2.4	1707	4	US-09-023-655-1060 Sequence 1060, Appl1
44	37	2.4	1192	4	US-09-583-447A-9 Sequence 9, Appl1
45	37	2.4	1633	4	US-09-583-447A-7 Sequence 7, Appl1

ALIGNMENTS

RESULT 1  
US-09-527-073-1  
Sequence 1, Application US/09527073  
Patent No. 6534313  
GENERAL INFORMATION:  
APPLICANT: Michael M. Neff  
APPLICANT: Joanne Chory  
TITLE OF INVENTION: GENETICALLY MODIFIED PLANTS HAVING  
TITLE OF INVENTION: MODULATED BRASSINOSTEROID SIGNALING  
FILE REFERENCE: SALKINS 024A  
CURRENT APPLICATION NUMBER: US/09/527, 073  
CURRENT FILING DATE: 2000-03-16  
PRIOR APPLICATION NUMBER: US 60/124570  
PRIOR FILING DATE: 1999-03-16  
PRIOR APPLICATION NUMBER: US 60/170,931  
PRIOR FILING DATE: 1999-12-14  
PRIOR APPLICATION NUMBER: US 60/172,832  
PRIOR FILING DATE: 1999-12-20  
NUMBER OF SEQ ID NOS: 16  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 1  
LENGTH: 1563  
TYPE: DNA  
ORGANISM: Arabidopsis thaliana  
FEATURE:  
OTHER INFORMATION: Oligonucleotide  
US-09-527-073-1

Query Match 100.0%; Score 1563; DB 4; Length 1563;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1563; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	ATGAGAGAAAGATGACAGCTGCTTATTCGAAAGTTTGTGCTGTAATCTTA	60
DB	1	ATGAGAGAAAGATGACAGCTGCTTATTCGAAAGTTTGTGCTGTAATCTTA	60
QY	61	AGCTCTTAATAGTGAAGGGATGCTCTGTAATGAGAGACCAAGAAAGATGAAGA	120
DB	61	AGCTCTTAATAGTGAAGGGATGCTCTGTAATGAGAGACCAAGAAAGATGAAGA	120
QY	121	CATTTCTTAACAAGAGATTCGAGGCTCTTATTTCTTCATCGAAATGTAAA	180
DB	121	CATTTCTTAACAAGAGATTCGAGGCTCTTATTTCTTCATCGAAATGTAAA	180
QY	181	GAACTGTGAAGATGATGCTTAAGCTTCTTCATCCATGCGCTTCTCAATATT	240
DB	181	GAACTGTGAAGATGATGCTTAAGCTTCTTCATCCATGCGCTTCTCAATATT	240
QY	241	CTTCTAGAGTCTCTCTTTTACCATCACTGAGAAATCTAAGGTGCTACATTTCTG	300



Db 241 CTTCCTAGAGTTCTCTCTTTTACATCATGAGAGAAAATCTACGGTCTACATTTCTG 300  
Qy 301 GTTGTGTCGGTCCAACTTCCGGTTAAACGTAGCCGATCCTGATTTGATCAGAGATC 360  
Db 301 GTTGTGTCGGTCCAACTTCCGGTTAAACGTAGCCGATCCTGATTTGATCAGAGATC 360  
Qy 361 TTCTTAAGTCTGAGTTCTACAGAGAAATGAAGTCAACCTTTGGTTAAACAATTGAA 420  
Db 361 TTCTTAAGTCTGAGTTCTACAGAGAAATGAAGTCAACCTTTGGTTAAACAATTGAA 420  
Qy 421 GGGGATGGACTACTAGTCTCAAGGTGAAAAATGGGCTCATCATCGAAAAATCATTAAC 480  
Db 421 GGGGATGGACTACTAGTCTCAAGGTGAAAAATGGGCTCATCATCGAAAAATCATTAAC 480  
Qy 481 CCTACTTTTCATATGAGAAATCTTAGTTGCTTACAGAGTTGTTGAAGTGTGACT 540  
Db 481 CCTACTTTTCATATGAGAAATCTTAGTTGCTTACAGAGTTGTTGAAGTGTGACT 540  
Qy 541 GATATGATGATTAATGTCCTCGATTAAGTTATCAGAAAACGGTGAAGTTAGGTATGTC 600  
Db 541 GATATGATGATTAATGTCCTCGATTAAGTTATCAGAAAACGGTGAAGTTAGGTATGTC 600  
Qy 601 TATGATGTTTCAAGTTTGAAGTGAAGTGTATTAAGTGAAGAGCTTTTGAAGTAC 660  
Db 601 TATGATGTTTCAAGTTTGAAGTGAAGTGTATTAAGTGAAGAGCTTTTGAAGTAC 660  
Qy 661 TATGAAGATGTCAGAGAGTTTTCAGTTCAGCTCAACAATGCTTTTGTCTGAA 720  
Db 661 TATGAAGATGTCAGAGAGTTTTCAGTTCAGCTCAACAATGCTTTTGTCTGAA 720  
Qy 721 GCTTTTCAAAAAGTCTTCAATCTCTGCTATTAAGTTTTCGCAACAGGGAATTGAG 780  
Db 721 GCTTTTCAAAAAGTCTTCAATCTCTGCTATTAAGTTTTCGCAACAGGGAATTGAG 780  
Qy 781 TCTCGAAGTTTACACAGAGATTAAGAGTGTGTTGAAGTCTATAGAGCGGCGAGA 840  
Db 781 TCTCGAAGTTTACACAGAGATTAAGAGTGTGTTGAAGTCTATAGAGCGGCGAGA 840  
Qy 841 CAAACGCTATATGATGAGAGAGGGAAGATGTAAGAGCCGGCGCGAAGATTTGTTG 900  
Db 841 CAAACGCTATATGATGAGAGAGGGAAGATGTAAGAGCCGGCGCGAAGATTTGTTG 900  
Qy 901 GGAATTAATGATCAGAGAAAGATGAGGTTTCAGGATTTGGAAGAGTGAAGAGC 960  
Db 901 GGAATTAATGATCAGAGAAAGATGAGGTTTCAGGATTTGGAAGAGTGAAGAGC 960  
Qy 961 TTTTCTTCCCGCGGAAACAGACAACTTCTAATCTGCTGAGCGTGAAGCACTTTGCTA 1020  
Db 961 TTTTCTTCCCGCGGAAACAGACAACTTCTAATCTGCTGAGCGTGAAGCACTTTGCTA 1020  
Qy 1021 TCCATGACCCCGAGTGGCAGGCAAGACGATGAGAGTCTCTAGAGTCTGCGGCTCA 1080  
Db 1021 TCCATGACCCCGAGTGGCAGGCAAGACGATGAGAGTCTCTAGAGTCTGCGGCTCA 1080  
Qy 1081 CGTATGTCCTTACCAAGAGCAATGCTTAAGCTTAAACGTTAGTATGATCTTGAAC 1140  
Db 1081 CGTATGTCCTTACCAAGAGCAATGCTTAAGCTTAAACGTTAGTATGATCTTGAAC 1140  
Qy 1141 GAGTCTTTAAAGTTGTTTCCACCAATGATGATGATGATGATGATGATGATGATG 1200  
Db 1141 GAGTCTTTAAAGTTGTTTCCACCAATGATGATGATGATGATGATGATGATGATG 1200  
Qy 1201 AAGCTAGAGAGGCTCAAAATCCCATGTCGACGAGACTTCAATCCCAATCAATAGCGTC 1260  
Db 1201 AAGCTAGAGAGGCTCAAAATCCCATGTCGACGAGACTTCAATCCCAATCAATAGCGTC 1260  
Qy 1261 CATCATGACCAAGCAATTTGGGGTAAATGAGAGTGAAGATTCATTCACCTGGTTGGC 1320  
Db 1261 CATCATGACCAAGCAATTTGGGGTAAATGAGAGTGAAGATTCATTCACCTGGTTGGC 1320  
Qy 1321 GATGAGTGGCGGCTGTCGCAAAACACCCGTTGCTTCAACGTTTGGCTCGAGTT 1380  
Db 1321 GATGAGTGGCGGCTGTCGCAAAACACCCGTTGCTTCAACGTTTGGCTCGAGTT 1380

Qy 1381 CGTACATGATTTGTCAGAAATCTTGTATATCTTACGCGCAAAATGACATCGCTATATG 1440  
Db 1381 CGTACATGATTTGTCAGAAATCTTGTATATCTTACGCGCAAAATGACATCGCTATATG 1440  
Qy 1441 ATCCACGCTTCACTTCTCACTTGGCTCTTACTTATGAGATGACCTACCGCTTATG 1500  
Db 1441 ATCCACGCTTCACTTCTCACTTGGCTCTTACTTATGAGATGACCTACCGCTTATG 1500  
Qy 1501 TTGCTTTATCTTCAACATGTCAGACCAATGACCTTCCGAGATTGACCAATCAAGAT 1560  
Db 1501 TTGCTTTATCTTCAACATGTCAGACCAATGACCTTCCGAGATTGACCAATCAAGAT 1560  
Qy 1561 TGA 1563  
Db 1561 TGA 1563

RESULT 2  
US-09-351-229-1  
; Sequence 1, Application US/09351229  
; Patent No. 6380465  
; GENERAL INFORMATION:  
; APPLICANT: Barlett, Michael  
; TITLE OF INVENTION: Cytochrome P450 Enzymes and Related Compounds and  
; FILE REFERENCE: P-1049  
; CURRENT APPLICATION NUMBER: US/09/351,229  
; EARLIER FILING DATE: 1999-07-11  
; EARLIER APPLICATION NUMBER: 60/092,596  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 1  
; LENGTH: 1584  
; TYPE: DNA  
; ORGANISM: Zee mays  
US-09-351-229-1

Query Match 17.3%; Score 270.4; DB 4; Length 1584;  
Best Local Similarity 51.5%; Pred. No. 1.1e-74; Indels 15; Gaps 2;  
Matches 680; Conservative 0; Mismatches 626;

Db 228 CTCTCAGATATTTCTTCTAGAGTCTCTCTTTTACATCATGAGAGAAATCTACGG 287  
Qy 228 CTCTCAGATATTTCTTCTAGAGTCTCTCTTTTACATCATGAGAGAAATCTACGG 287  
Db 249 CTGCCAGACATCACCCACGCGTGCAGCCATGATCAGACACATCAAGAAATCGG 308  
Qy 249 CTGCCAGACATCACCCACGCGTGCAGCCATGATCAGACACATCAAGAAATCGG 308  
Db 288 TGCTACATTTCTGTTGGTTGGTCCAACTTCCGGTTAAACGGTACCTGATTT 347  
Qy 288 TGCTACATTTCTGTTGGTTGGTCCAACTTCCGGTTAAACGGTACCTGATTT 347  
Db 309 GAAACTATGCTTACCTGCTGGTGGCCAAACCAAGGATGATCTTGAACCAAGATT 368  
Qy 309 GAAACTATGCTTACCTGCTGGTGGCCAAACCAAGGATGATCTTGAACCAAGATT 368  
Db 348 GATCAGAGAGATCTTCTTAAGTCTGAGTTCTACAGAGATGAAGTCAACCTTTGTT 407  
Qy 348 GATCAGAGAGATCTTCTTAAGTCTGAGTTCTACAGAGATGAAGTCAACCTTTGTT 407  
Db 369 AGTCAGAGAGTCTCTCTTAAGTCTGAGTTCTACAGAGATGAAGTCAACCTTTGTT 428  
Qy 369 AGTCAGAGAGTCTCTCTTAAGTCTGAGTTCTACAGAGATGAAGTCAACCTTTGTT 428  
Db 408 TAAACACTTGAAGGCGATGACTTCTTAAGTCTTCAAGAGTGAAGAAATGGGCTCATACG 467  
Qy 408 TAAACACTTGAAGGCGATGACTTCTTAAGTCTTCAAGAGTGAAGAAATGGGCTCATACG 467  
Db 429 TGGAGAGTGTAGCCCAAGCGGCTTGTAAATGATGATGATGATGATGATGATGATG 488  
Qy 429 TGGAGAGTGTAGCCCAAGCGGCTTGTAAATGATGATGATGATGATGATGATGATG 488  
Db 468 AAAAATCATTAAGCCCTACTTTTCATATGAGAGATCTTAAGTGTCTTACAGAGTTGTT 527  
Qy 468 AAAAATCATTAAGCCCTACTTTTCATATGAGAGATCTTAAGTGTCTTACAGAGTTGTT 527  
Db 489 GAGATTTCTTAATCTGCACTTTCACAGAGAAATTAAGGAGATGATGCCAGATTTTC 548  
Qy 489 GAGATTTCTTAATCTGCACTTTCACAGAGAAATTAAGGAGATGATGCCAGATTTTC 548  
Db 528 GAAAGTGTGACTGATATGATGATGATGATGATGATGATGATGATGATGATGATG 587  
Qy 528 GAAAGTGTGACTGATATGATGATGATGATGATGATGATGATGATGATGATGATG 587  
Db 549 TACCTGCTGATTAATGATTAATTAATGATGATGATGATGATGATGATGATGATG 608  
Qy 549 TACCTGCTGATTAATGATTAATTAATGATGATGATGATGATGATGATGATGATG 608  
Db 588 TGAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 647  
Qy 588 TGAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 647  
Db 609 TGAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 668  
Qy 609 TGAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 668  
Db 648 TTTTGAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 707  
Qy 648 TTTTGAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 707

Db 669 GTTTGGAGCAACTATCAAGAGGAGAGATTTTGTAGCTACAGAGAACTAGCTGA 728  
 QY 708 TCTTTGTGCTGAAGCTTTTCAAAAAGCTTCACTTCCTGCTATGATTTTTCGACAG 767  
 Db 729 AGCCTCAATCCAACTGTGTAGACAAATTTATCCAGGCTATGTGTTCTTGCCACAA 788  
 QY 768 AGGGAATTTGAAGTCCGAAGTGAAGAGAGATGAAGAGTCTGTGTTGAAGCTGAT 827  
 Db 789 AAACAACAGAGAGATGAAGCAATGATGTAGAGATCCGCAAAATTCCTCGTGAATAT 848  
 QY 828 AAGCGCGGAGAGCAAAACGCTATATGATGAGAGAGGAGAAATGTAAAGG----- 879  
 Db 849 TGGAGAGAGAGAGAGATTAATAAAACAGAGAAACAAATAAAGTACTGTGCGGCTT 908  
 QY 880 ----CCGGGGGAGAGATTTGTGGATTAATGATGAGCAAAAGATGACGCTTCA 935  
 Db 909 ATTACTGAGTCAAAACAGAGCAATCAATGAAATCAAGCTCGGATTTGACAAAGA 968  
 QY 936 GGCATTTGAGAGAGTGAAGCTTTTCTTCGCGGAGAAACAGCACTTCTAATCT 995  
 Db 969 AGATGTGATTTGAGATGCAATTTTACTTTGAGATTTGAGACATCAGTCTT 1028  
 QY 996 GCTGACGTGAGACCACTTTGTATCCATGACCCGAGAGTGGCCAGGCAAGACGTGA 1055  
 Db 1029 GCTTACTTGGACACTTATGTGTCTAGACCCCAAGATGGCAAGAGAGAGAGAGA 1088  
 QY 1056 TGAAGTCTCAGAGGTGTGGGCTCAGTATGCTCCCTACCAAGACATGCTGTTAAGCT 1115  
 Db 1089 AGAGGTTTGAAGCACTTTGG--AAGAACACACCAATATATGATTTGAGACCGCTT 1145  
 QY 1116 TAAACGTTGAGTATGATCTTGAACGAGCTTTTAAAGTTGATTCACCAATATGATGCTAC 1175  
 Db 1146 CAAAGCTGTAACATGATTTCTACATGAGGCTCCTTAGGTTGATCCACCGGCAACCTTCT 1205  
 QY 1176 GATTCAGACGCGTAAATCGGATGTGAAGGTAGAGGCTTAAATCCCATGTGGACGCA 1235  
 Db 1206 AACAGAGAACTTAATAGAGAAATGAGCTCGGTGAATCAAAATCTTCGACAGAGTGA 1255  
 QY 1236 GCTTCTAATCCCAATCAATAGCGGTCCATCATGACCAAGCCATTTGGGGTAATGACGTGA 1295  
 Db 1266 GCTCCTTCTCCCGTCACTTCAATCAATGATCCGACATTTGGGGGAAAAAGCGCAAG 1325  
 QY 1296 CGAATTCATTCAGCTCGTTTTCGAGTGAAGTGGCGGCTGTGCAAAACCCCGCTGG 1355  
 Db 1326 CGAGTTCACCCAGAAAGGTTTCCAAACGCAATCTTCACGCAACCGAGCATGAGCTGC 1385  
 QY 1356 CTTCAATACGTTGGGCTCGAGTGTGTAATGATGATGATGATGATGATGATGATGATGAT 1415  
 Db 1386 TTTCTTCCGTTGGAGAGGGGCCCCAGAGATCTGCAATCGGCGAGAGGTTTGGCTTCTG 1445  
 QY 1416 GGGCAATTAACACTGCTGTATATGATTCACGCTTCACTTCACTTCACTTCACTTCA 1475  
 Db 1446 AGCCAGATGACCTATGACACATCTTCAGCGCTTCTGTTGAGCTTCACTTCACTTCA 1505  
 QY 1476 TCGAGATGACCTACCTGCTTATGATGATGATGATGATGATGATGATGATGATGATGAT 1535  
 Db 1506 CACCCAGCGGCGGTAACCTGTATACACTGACCTTCAAGCAGGTCCTCAATAGGCT 1565  
 QY 1536 C 1536  
 Db 1566 C 1566

RESULT 3  
 US-09-527-073-16  
 ; Sequence 16, Application US/09527073  
 ; Patent No. 6534313  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Michael M. Neff  
 ; APPLICANT: Joanne Chory  
 ; TITLE OF INVENTION: GENETICALLY MODIFIED PLANTS HAVING  
 ; TITLE OF INVENTION: MODULATED BRASSINOSTEROID SIGNALING  
 ; FILE REFERENCE: SALKINS\_024A

; CURRENT APPLICATION NUMBER: US/09/527,073  
 ; CURRENT FILING DATE: 2000-03-16  
 ; PRIOR APPLICATION NUMBER: US 60/124570  
 ; PRIOR FILING DATE: 1999-03-16  
 ; PRIOR APPLICATION NUMBER: US 60/170,931  
 ; PRIOR FILING DATE: 1999-12-14  
 ; PRIOR APPLICATION NUMBER: US 60/172,832  
 ; PRIOR FILING DATE: 1999-12-20  
 ; NUMBER OF SEQ ID NOS: 16  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 16  
 ; LENGTH: 588  
 ; TYPE: DNA  
 ; ORGANISM: Arabidopsis thaliana  
 ; FEATURE:  
 ; NAME/KEY: misc.feature  
 ; LOCATION: (1)...(588)  
 ; OTHER INFORMATION: n = A,T,C or G  
 US-09-527-073-16  
 Query Match 11.8%; Score 184; DB 4; Length 588;  
 Best Local Similarity 99.0%; Pred. No. 1e-47;  
 Matches 206; Conservative 0; Mismatches 0; Indels 2; Gaps 2;

QY 1 ATGAGAGAAAGATGAGAGCTGTTCACTTCGAAAGTTCTGTCTGTATCTTA 60  
 Db 382 ATGAGAGAAAGATGAGAGCTGTTCACTTCGAAAGTTCTGTCTGTATCTTA 441  
 QY 61 AGT-CTTGAATATGAGAGGATGCTCTGTATGTTGAGACCAAGAAATTTAGA 115  
 Db 442 AGTCTGTGAATGAGAGGATGCTCTGTATGTTGAGACCAAGAAATTTAGA 501  
 QY 120 ACATTTCTTAAACAGAAATTCGAGTCTCTTATCTTCAATCGAAATGTTAA 179  
 Db 502 ACATTTCTTAAACAGAAATTCGAGTCTCTTATCTTCAATCGAAATGTTAA 561  
 QY 180 AGAATCTTGTGAATGATGCTTAAAGCT 207  
 Db 562 AGAATCTTGT-GAATGATGCTTAAAGCT 588

RESULT 4  
 US-09-615-192A-101  
 ; Sequence 101, Application US/09615192A  
 ; Patent No. 6410718  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Blockberg, Leonard N.  
 ; APPLICANT: Havukkala, Ilkka  
 ; TITLE OF INVENTION: Materials and Methods for the  
 ; TITLE OF INVENTION: Modification of Plant Lipid Content  
 ; FILE REFERENCE: 11000.1003c4U  
 ; CURRENT APPLICATION NUMBER: US/09/615,192A  
 ; CURRENT FILING DATE: 2000-07-12  
 ; PRIOR APPLICATION NUMBER: US 08/975,316  
 ; PRIOR FILING DATE: 1997-11-21  
 ; PRIOR APPLICATION NUMBER: US 08/713,000  
 ; PRIOR FILING DATE: 1996-09-11  
 ; PRIOR APPLICATION NUMBER: US 09/169,789  
 ; PRIOR FILING DATE: 1998-10-09  
 ; NUMBER OF SEQ ID NOS: 405  
 ; SOFTWARE: FastSeq for Windows Version 3.0  
 ; SEQ ID NO 101  
 ; LENGTH: 612  
 ; TYPE: DNA  
 ; ORGANISM: Eucalyptus grandis  
 US-09-615-192A-101

Query Match 4.8%; Score 75.8; DB 4; Length 612;  
 Best Local Similarity 50.1%; Pred. No. 1.2e-13;  
 Matches 216; Conservative 0; Mismatches 212; Indels 3; Gaps 1;  
 QY 75 GAAGGATGCTCTCTGTATGTTGAGAGACCAAGAAATTTGAAGAACTTTCTTAACA 134

Db 185 GAGGCGGTGAACTGGGTGTGCTGAGGCCGAGAGGCTCGAGAGGCTTCTGAGACAGCA 244  
 Qy 135 AGGAAATTCGAGGTCTCTCTTATCATTTCTTCATCGGAAATGTTAAGAACTTGTGGAA 194  
 Db 245 AGGTCTCTCCGCGAGTCTCAACCTTCTGTGTGGGCACTCGAAGGAACTTGGCGAT 304  
 Qy 195 GATGCTTAAAGCTTCTTCTCATCTCATGCTTCTCTCATCAATTTCTCTTCAAGTCT 254  
 Db 305 GCTCAAGGAGGCGCAATCCAGCCGCTCTCCGATGACATCAAGCTCTCTCTT 354  
 Qy 255 CTCTTTTACCATCTGAGAGAAATCTACGCTGCTCATTTCTGTGTTGTGCTGCTC 314  
 Db 365 GCTTTCTTGATCATCTTCTTCAACCTTGTGAGAAACCTGTTCAATGAGATGGGCC 424  
 Qy 315 AACTTTCGGTTAAGCGTAGCCGATCTGATTTGATCGAGAGATCTTCTTAAGTCTGA 374  
 Db 425 AACACCAAGAGTGAATATACGACCCGAGACCAATTAAGAGGATTTCTTAAGATATA 484  
 Qy 375 GTTCTACGAGAAATGAGAGCTCAACCTTTGTTTAAACAATGAGGCGATGAGTACT 434  
 Db 485 TGAATATCCCAAGCCGCTCTCAATCCCTGTGAAAGTGTCT--CGTGAATGACTCGC 541  
 Qy 435 TAGTCTCAAGGTGAAAAATGAGCTCATCATGAAAAATCATTTAGCCCTTACTTTTCA 494  
 Db 542 GAACCATGAGGGGAGAAATGGGCTCGGACCGAAAGATTTATCATTCAGATTCACAT 601  
 Qy 495 GAGAAATCTTA 505  
 Db 602 GGAGAAATGTA 612

RESULT 5

US-09-169-789-101

; Sequence 101, Application US/09169789  
 ; Patent No. 6635328  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Bloksberg, Leonard N.  
 ; APPLICANT: Havakkala, Ilkka  
 ; TITLE OF INVENTION: Materials and Methods for the  
 ; FILE REFERENCE: 11000.1003c2  
 ; CURRENT APPLICATION NUMBER: US/09/169,789  
 ; CURRENT FILING DATE: 1998-10-09  
 ; EARLIER APPLICATION NUMBER: US 08/975,316  
 ; EARLIER FILING DATE: 1997-11-21  
 ; EARLIER APPLICATION NUMBER: US 08/713,000  
 ; EARLIER FILING DATE: 1996-09-11  
 ; NUMBER OF SEQ ID NOS: 195  
 ; SOFTWARE: FastSeq for Windows Version 3.0  
 ; SEQ ID NO 101  
 ; LENGTH: 612  
 ; TYPE: DNA  
 ; ORGANISM: Eucalyptus grandis  
 US-09-169-789-101

Query Match 4.8%; Score 75.8; DB 4; Length 612;

Best Local Similarity 50.1%; Pred. No. 1.2e-13;  
 Matches 216; Conservative 0; Mismatches 212; Indels 3; Gaps 1;

Qy 75 GAAGGATATGCTCTGTTATGTGAGACCAAGAAAGATTAAGAACTTCTTAACA 134  
 Db 185 GAGGCGGTGAACTGGGTGTGCTGAGGCCGAGAGGCTCGAAGGCTTCTGAGACAGA 244  
 Qy 135 AGAATTCGAGGTCTCTCTTATCATTTCTTCATCGGAAATGTTAAGAACTTGTGGAA 194  
 Db 245 AGGTCTCTCCGCGAGTCTCAACCTTCTGTGTGGGCACTCGAAGGAACTTGGCGAT 304  
 Qy 195 GATGCTTAAAGCTTCTTCTCATCTCATGCTTCTCTCATCAATTTCTCTTCAAGTCT 254  
 Db 305 GCTCAAGGAGGCGCAATCCAGCCGCTCTCCGATGACATCAAGCTCTCTCTT 354  
 Qy 255 CTCTTTTACCATCTGAGAGAAATCTACGCTGCTCATTTCTGTGTTGTGCTGCTC 314

Db 365 GCTTTCTTGATCATCTTCTTCAACCTTATGGAAGAGCTCGTTCATGATGGGCC 424  
 Qy 315 AACTTTCGGTTAAGCGTAGCCGATCTGATTTGATCGAGAGATCTTCTTAAGTCTGA 374  
 Db 425 AACACCAAGAGTGAATATACGACCCGAGACCAATTAAGAGGATTTCTTAAGATATA 484  
 Qy 375 GTTCTACGAGAAATGAGAGCTCAACCTTTGTTTAAACAATGAGGCGATGAGTACT 434  
 Db 485 TGAATATCCCAAGCCGCTCTCAATCCCTGTGAAAGTGTCT--CGTGAATGACTCGC 541  
 Qy 435 TAGTCTCAAGGTGAAAAATGAGCTCATCATGAAAAATCATTTAGCCCTTACTTTTCA 494  
 Db 542 GAACCATGAGGGGAGAAATGGGCTCGGACCGAAAGATTTATCATTCAGATTCACAT 601  
 Qy 495 GAGAAATCTTA 505  
 Db 602 GGAGAAATGTA 612

RESULT 6

US-09-615-192A-169

; Sequence 169, Application US/09615192A  
 ; Patent No. 6410718  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Bloksberg, Leonard N.  
 ; APPLICANT: Havakkala, Ilkka  
 ; TITLE OF INVENTION: Materials and Methods for the  
 ; FILE REFERENCE: 11000.1003c4U  
 ; CURRENT APPLICATION NUMBER: US/09/615,192A  
 ; CURRENT FILING DATE: 2000-07-12  
 ; PRIOR APPLICATION NUMBER: US 08/975,316  
 ; PRIOR FILING DATE: 1997-11-21  
 ; PRIOR APPLICATION NUMBER: US 08/713,000  
 ; PRIOR FILING DATE: 1996-09-11  
 ; PRIOR APPLICATION NUMBER: US 09/169,789  
 ; PRIOR FILING DATE: 1998-10-09  
 ; NUMBER OF SEQ ID NOS: 405  
 ; SOFTWARE: FastSeq for Windows Version 3.0  
 ; SEQ ID NO 169  
 ; LENGTH: 398  
 ; TYPE: DNA  
 ; ORGANISM: Eucalyptus grandis  
 US-09-615-192A-169

Query Match 4.7%; Score 73.4; DB 4; Length 398;

Best Local Similarity 51.2%; Pred. No. 5.3e-13;  
 Matches 198; Conservative 0; Mismatches 186; Indels 3; Gaps 1;

Qy 563 ATAAATTCAGAAAAAGGCTGAAGTTGAGTATGATGCTTATGAGTGTTCAGATTTTGA 622  
 Db 6 ATTTGATATCGAGAGAGATCTCTGAGATGAGAGCGGTGGGTGACCTTCAAAATTTGA 65  
 Qy 623 CTGAAGATGTTATTAAGTGAACAGCTTTTGAAGTATGAGTGAAGATGAGAGT 682  
 Db 66 CCGGTAGAGTGAATCTCTGACAGCGTGTGGAGTAGCTTCGAAAGGCAAGAGATCT 125  
 Qy 683 TTGAGCTTCAAGCTCAACAATGCTTCTTGTGCTGAAGCTTTTCAAAAGCTTCATTC 742  
 Db 126 CCGAATCTTCAAGGGGAAACAAGCCAGCTACATATAATGCGCTTCAATCGGCTCATTC 185  
 Qy 743 CTGCTATATGATTTTTCGACAAAGGGAATTTGAAGTCTCGGAAGTTAGCAAGAGA 802  
 Db 186 CTGTTGAGAGTTTGTGCAACTTAAGANTAAAGAGAGTGAAGAGCAATAGTAAAGA-- 243  
 Qy 803 TTAGGAAGTGTGTTGAAGCTGATAGAGCGGCGAGACCAAAACCTTATAGTAGAGAAG 862  
 Db 244 -AGTGGGAGTCTGCTCTCAATGACATATCCGAGAAAGAGAAAGCAATTAAGGAAAGGG 302  
 Qy 863 GGGAGAAATTAAGAGCCGCGCGAGAGATTTGTTGGATTTAATGATTCAGCAAGA 922  
 Db 303 AAGCTGCTGCGATGATCTGCTGGGCTGTGCTGAGTCAAAACATGAAGGAATGTGC 362



Mon May 24 08:10:08 2004

us-09-992-901-1.rn1

Page 6

TOPOLOGY: linear  
IMMEDIATE SOURCE:  
CLONE: pTZ9pt-Fls  
US-08-232-463-14

Query Match 3.4%; Score 53; DB 1; Length 7218;  
Best Local Similarity 9.3%; Pred. No. 8.9e-06;  
Matches 23; Conservative 137; Mismatches 87; Indels 0; Gaps 0;

Qy 762 GACAGAGGGAATTGAGCTGAGGATTAGACAGAGGAGATGAGGATGCTGTGATAA 821  
Db 1262 RRR 1203  
Qy 822 GCTGATAGAGCGGCGAGACAAACGCTATGATGAGAGGAGAGATGTAAGAGCC 881  
Db 1202 RRR 1143  
Qy 882 GCGCGGAGAGATTGTTGGGATTATGATTCAGCAAGATGACGCTTACAGACAT 941  
Db 1142 RRR 1083  
Qy 942 TGTGAGAGAGTGAAGCTTTTCTTCGCGGGAACAGACACTTCTAATCTGCTGAC 1001  
Db 1082 RRR 1023  
Qy 1002 GTGAGCG 1008  
Db 1022 GTGAGCG 1016

RESULT 10  
US-08-118-554-55  
Sequence 55, Application US/09118554A  
Patent No. 6365348  
GENERAL INFORMATION:  
APPLICANT: Reed, Steven G.  
APPLICANT: Xu, Jiangchun  
TITLE OF INVENTION: COMPOUNDS FOR DIAGNOSIS OF BREAST CANCER AND  
FILE REFERENCE: 210121.450C1  
CURRENT APPLICATION NUMBER: US/09/118,554A  
CURRENT FILING DATE: 1998-07-17  
EARLIER APPLICATION NUMBER: 08/988,255  
EARLIER FILING DATE: 1997-12-24  
NUMBER OF SEQ ID NOS: 67  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 55  
LENGTH: 933  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-09-118-554-55

Query Match 3.1%; Score 48; DB 4; Length 933;  
Best Local Similarity 64.3%; Pred. No. 9.1e-05;  
Matches 72; Conservative 0; Mismatches 40; Indels 0; Gaps 0;

Qy 1342 AAACACCCCGTTGGCTTCATACCGTTGGCGCTGAGATTGATCATGATGATGATGAT 1401  
Db 329 ATACATCCCTATGCTTCATACCATTTCTCAGCTGATTAAGAGATGATGATGATGATGAT 388  
Qy 1402 CTGCTATACTTACGAGCAATTAAGCACTGCTGTAATGATTCACAGCTTCA 1453  
Db 389 TTGGCATTAATGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 440

RESULT 11  
US-09-118-627-55  
Sequence 55, Application US/09118627A  
Patent No. 6379951  
GENERAL INFORMATION:  
APPLICANT: Reed, Steven G.  
APPLICANT: Xu, Jiangchun  
TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY OF BREAST CANCER

TITLE OF INVENTION: AND METHODS FOR THEIR USE  
FILE REFERENCE: 210121.446C1  
CURRENT APPLICATION NUMBER: US/09/118,627A  
CURRENT FILING DATE: 1998-07-17  
NUMBER OF SEQ ID NOS: 67  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 55  
LENGTH: 933  
TYPE: DNA  
ORGANISM: Homo sapien  
US-09-118-627-55

Query Match 3.1%; Score 48; DB 4; Length 933;  
Best Local Similarity 64.3%; Pred. No. 9.1e-05;  
Matches 72; Conservative 0; Mismatches 40; Indels 0; Gaps 0;

Qy 1342 AAACACCCCGTTGGCTTCATACCGTTGGCGCTGAGATTGATCATGATGATGATGAT 1401  
Db 329 ATACATCCCTATGCTTCATACCATTTCTCAGCTGATTAAGAGATGATGATGATGATGAT 388  
Qy 1402 CTGCTATACTTACGAGCAATTAAGCACTGCTGTAATGATTCACAGCTTCA 1453  
Db 389 TTGGCATTAATGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 440

RESULT 12  
US-09-602-877A-55  
Sequence 55, Application US/09602877A  
Patent No. 6432707  
GENERAL INFORMATION:  
APPLICANT: Reed, Steven G.  
APPLICANT: Xu, Jiangchun  
APPLICANT: Dillon, David C.  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
FILE REFERENCE: 210121.446C5  
CURRENT APPLICATION NUMBER: US/09/602,877A  
CURRENT FILING DATE: 2000-06-22  
NUMBER OF SEQ ID NOS: 107  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 55  
LENGTH: 933  
TYPE: DNA  
ORGANISM: Homo sapien  
US-09-602-877A-55

Query Match 3.1%; Score 48; DB 4; Length 933;  
Best Local Similarity 64.3%; Pred. No. 9.1e-05;  
Matches 72; Conservative 0; Mismatches 40; Indels 0; Gaps 0;

Qy 1342 AAACACCCCGTTGGCTTCATACCGTTGGCGCTGAGATTGATCATGATGATGATGAT 1401  
Db 329 ATACATCCCTATGCTTCATACCATTTCTCAGCTGATTAAGAGATGATGATGATGATGAT 388  
Qy 1402 CTGCTATACTTACGAGCAATTAAGCACTGCTGTAATGATTCACAGCTTCA 1453  
Db 389 TTGGCATTAATGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 440

RESULT 13  
US-09-602-877A-102  
Sequence 102, Application US/09602877A  
Patent No. 6432707  
GENERAL INFORMATION:  
APPLICANT: Reed, Steven G.  
APPLICANT: Xu, Jiangchun  
APPLICANT: Dillon, David C.  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
FILE REFERENCE: 210121.446C5  
CURRENT APPLICATION NUMBER: US/09/602,877A  
CURRENT FILING DATE: 2000-06-22  
NUMBER OF SEQ ID NOS: 107

Mon May 24 08:10:08 2004

US-09-992-901-1.rnt

Page 7

SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 102  
; LENGTH: 940  
; TYPE: DNA  
; ORGANISM: Human  
US-09-602-877A-102

Query Match 3.1%; Score 48; DB 4; Length 940;  
Best Local Similarity 64.3%; Pred. No. 9.2e-05;  
Matches 72; Conservative 0; Mismatches 40; Indels 0; Gaps 0;

QY 1342 AACACCCCGTGGTTCATACCGTTGGCTCGAGTGTGATCATGTCAGAT 1401  
DB 330 ATACATCCCTATGCTTCATACCATTTTCAGCTGATTAAGSACTGCGAGCAT 389  
QY 1402 CTGGTATACCTCAGGCCAATTGACACTGCTGTAAATGATCCAGCTTCA 1453  
DB 390 TTGGCATATATGAGTGAAGTGCGAGTGCATTAACTGCTCCGCTCA 441

RESULT 14  
US-08-991-677-1  
; Sequence 1, Application US/08991677A  
; Patent No. 6252135  
; GENERAL INFORMATION:  
; APPLICANT: Chiang, Vincent L  
; APPLICANT: Caraway, Daniel T  
; APPLICANT: Smeltzer, Richard H  
; TITLE OF INVENTION: Production of Syringyl Lignin in Gymnosperms  
; FILE REFERENCE: 50617  
; CURRENT APPLICATION NUMBER: US/08/991,677A  
; CURRENT FILING DATE: 1997-12-16  
; EARLIER APPLICATION NUMBER: US 60/033,381  
; EARLIER FILING DATE: 1996-12-16  
; NUMBER OF SEQ ID NOS: 11  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 1  
; LENGTH: 1708  
; TYPE: DNA  
; ORGANISM: Liquidambar styraciflua  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (48) .. (1571)  
US-08-991-677-1

Query Match 3.0%; Score 46.2; DB 3; Length 1708;  
Best Local Similarity 46.2%; Pred. No. 0.0049;  
Matches 227; Conservative 0; Mismatches 258; Indels 6; Gaps 2;

QY 965 TCTTCGCCGGAACAACACACTTCTAATCTGCTGACGTCGACGACATCTTGTATCA 1024  
DB 937 TCATCTCAGGATGACACACCGCAATCTCTGTAATGGGCCATGGCCAGTTAATTA 996  
QY 1025 TGAACCCGAGTGGCGGCCAAGCAGTGAAGTGAAGTCTCAAGGTCGCGCTCAGTG 1084  
DB 997 AGAACCAGAGGTGACAAAGGCTCAAGAGACTGACATATCTTGGGTCGAGC 1056  
QY 1085 ATGTCCCTACCAAGACATGTCGTTAAGCTTAAACGTTGAGTGAATCTTGAACAGT 1144  
DB 1057 GTGTCTGACCGAATTTGACTCTCAAGCTCCCTTATCTAATAGTGTACCAAGGAG 1116  
QY 1145 CTTTAAGGTTGTATCCACCAATAGTGTACAT---TCGAGCGGCTAAATCGAGTGA 1201  
DB 1117 CACTAAGGCTGACCTCCCAACACCTAATGCTCTCTATGCGCCATGCAAGTCA 1176  
QY 1202 AGCTAGAGGATCAAAATCCCATGTGACGAGCTTCTAATCCCAATCAAGGCTCC 1261  
DB 1177 AATTTGGTGTCTGACATCTTCAAGGATCAAAATGTTATGTAATGTCTGGGCGGTG 1236  
QY 1262 ATATGACCAAGCACTTTGGGATATGAGTGAAGAAATTCATCACTGCTGTTGGG 1321  
DB 1237 CTCGTATCCAGGATGTGTGGC--TGACCACATAGATTTCGACCGGAAGGTTCTGTG 1293

QY 1322 ATGAGTGGCGCGGCTGTCGCAACACCCGTTGCTTCAATACGTTGGCTCGAGTTC 1381  
DB 1294 AAGACATGTGACATATAAAGGTGACATTATAGGCTACTGCGCTTGGTGGAGGAGC 1353  
QY 1382 GTACATGATTTGTGACAAATCTTGTCTAATCTTCAAGGCCAATTGACATCTGTATGA 1441  
DB 1354 GTGTTGGCCCGGTGACAACTTGATTCATTTGGTCAATGATGATGGTCACTAT 1413  
QY 1442 TCCAGCGTTC 1452  
DB 1414 TGACCATTTTC 1424

RESULT 15  
US-09-103-840A-2/c  
; Sequence 2, Application US/09103840A  
; Patent No. 6294328  
; GENERAL INFORMATION:  
; APPLICANT: FLEISCHMAN, Robert D.  
; APPLICANT: WHITE, Owen R.  
; APPLICANT: FRASER, Claire M.  
; APPLICANT: VENTER, John C.  
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM  
; FILE REFERENCE: 24366-20007.00  
; CURRENT APPLICATION NUMBER: US/09/103,840A  
; CURRENT FILING DATE: 1998-06-24  
; NUMBER OF SEQ ID NOS: 2  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 2  
; LENGTH: 4403765  
; TYPE: DNA  
; ORGANISM: Mycobacterium tuberculosis  
; FEATURE:  
; OTHER INFORMATION: CDC 1551  
; OTHER INFORMATION: "r" bases at various positions throughout the sequence  
; OTHER INFORMATION: represent a, t, c or g  
US-09-103-840A-2

Query Match 2.9%; Score 45.8; DB 3; Length 4403765;  
Best Local Similarity 53.0%; Pred. No. 0.097;  
Matches 98; Conservative 0; Mismatches 87; Indels 0; Gaps 0;

QY 1357 TTCAATCCGTTGGCTGAGTTCGTACATGATGATGATCAACGCTTCACTTCACTTGGCTCTACTTAT 1416  
DB 1569612 TACCTGCCGCTTGGTGGGCGCGACGATCTGATCGAGACGACTTCCGCTGATGAG 1569553  
QY 1417 GCCAATGACACTGCTGTATGATGATCAACGCTTCACTTCACTTGGCTCTACTTAT 1476  
DB 1569552 AAGCTTTGATGGCGGCAATATGATGATCAACTTCACTTGACTCGCCCTGGCTAT 1569493  
QY 1477 CAGCATGACCTACCGTCTTATGTTATCTTATCTTAATCATGTGTGACCAATACCTTC 1536  
DB 1569492 CAGCTGAACTCGAGCGGACGCTGACCTGCGACCAAGCAGAGTACGTCATCGG 1569433  
QY 1537 CGAG 1541  
DB 1569432 AGGAG 1569428

Search completed: May 22, 2004, 20:15:52  
Job time : 157 secs

Mon May 24 08:10:08 2004

us-09-992-901-1.rnpb

Page 1

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using SW model

Run on: May 22, 2004, 19:00:57 ; Search time 730 Seconds

(without alignments)  
9730.698 Million cpi) updates/sec

Title: US-09-992-901-1

Perfect score: 1563

Sequence: 1 atgagagaagaagaatgacag.....tgaccatcatgagattga 1563

Scoring table: IDENTITY NUC

Searched: 2953838 seqs, 2272363821 residues

Total number of hits satisfying chosen parameters: 5907676

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Database:

Published Applications NA:  
1: /cgn2\_6/ptodata/1/pubna/US07\_PUBCOMB.seq:  
2: /cgn2\_6/ptodata/1/pubna/PCR\_NEW\_PUB.seq:  
3: /cgn2\_6/ptodata/1/pubna/US06\_NEW\_PUB.seq:  
4: /cgn2\_6/ptodata/1/pubna/US06\_PUBCOMB.seq:  
5: /cgn2\_6/ptodata/1/pubna/US07\_NEW\_PUB.seq:  
6: /cgn2\_6/ptodata/1/pubna/PCR\_NEW\_PUBCOMB.seq:  
7: /cgn2\_6/ptodata/1/pubna/US08\_NEW\_PUB.seq:  
8: /cgn2\_6/ptodata/1/pubna/US08\_PUBCOMB.seq:  
9: /cgn2\_6/ptodata/1/pubna/US09\_PUBCOMB.seq:  
10: /cgn2\_6/ptodata/1/pubna/US09B\_PUBCOMB.seq:  
11: /cgn2\_6/ptodata/1/pubna/US09C\_PUBCOMB.seq:  
12: /cgn2\_6/ptodata/1/pubna/US09\_NEW\_PUB.seq:  
13: /cgn2\_6/ptodata/1/pubna/US09\_NEW\_PUBCOMB.seq:  
14: /cgn2\_6/ptodata/1/pubna/US10B\_PUBCOMB.seq:  
15: /cgn2\_6/ptodata/1/pubna/US10C\_PUBCOMB.seq:  
16: /cgn2\_6/ptodata/1/pubna/US10C\_PUBCOMB.seq:  
17: /cgn2\_6/ptodata/1/pubna/US10C\_PUBCOMB.seq:  
18: /cgn2\_6/ptodata/1/pubna/US60\_NEW\_PUB.seq:  
19: /cgn2\_6/ptodata/1/pubna/US60\_PUBCOMB.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1563	100.0	1563	9	US-09-992-901-1
2	493.2	31.6	1236	13	US-10-424-599-59416
3	303	19.4	1778	13	US-10-424-599-40486
4	291.4	18.6	1844	13	US-10-424-599-56717
5	274.8	17.6	1886	13	US-10-425-114-30156
6	273.2	17.5	1836	13	US-10-425-114-22333
7	273.2	17.5	1877	13	US-10-425-114-13336
8	267.8	17.1	1853	13	US-10-425-114-10433
9	254.2	16.3	1922	13	US-10-424-599-85977
10	250.8	16.0	1872	9	US-10-424-599-126502
11	241.2	15.4	1557	13	US-09-938-842A-92
12	241.2	15.4	1574	11	US-09-938-842A-92
13	232.4	14.9	1574	13	US-10-425-114-15634
14	231.4	14.8	4300	13	US-10-424-599-74806

15	230.2	14.7	1808	13	US-10-425-114-24276	Sequence 24276, A
16	218.6	14.0	1704	6	US-10-429-949-22	Sequence 22, Appl
17	194.8	12.5	1539	9	US-09-938-842A-379	Sequence 379, Appl
18	194.8	12.5	1539	11	US-09-938-842A-379	Sequence 379, Appl
19	191.2	12.2	1688	13	US-10-425-114-476	Sequence 476, Appl
20	185.6	11.9	1747	13	US-10-425-114-23110	Sequence 23110, A
21	184	11.8	588	9	US-09-992-901-16	Sequence 16, Appl
22	184	11.8	1786	13	US-10-425-114-13600	Sequence 13600, A
23	184	11.5	1781	13	US-10-425-114-26722	Sequence 26722, A
24	179.4	11.5	1367	13	US-10-424-599-104525	Sequence 104525, A
25	173.2	11.1	1539	13	US-10-425-114-28233	Sequence 2823, Appl
26	172.8	11.1	1126	13	US-10-425-114-28252	Sequence 28252, Appl
27	170.8	10.9	2188	13	US-10-424-599-76177	Sequence 76177, A
28	168.4	10.8	522	13	US-10-424-599-93836	Sequence 93836, A
29	160.2	10.2	1843	13	US-10-424-599-54624	Sequence 54624, A
30	153.6	9.8	1124	9	US-09-970-445-88	Sequence 88, Appl
31	147.6	9.4	1789	13	US-10-425-114-28468	Sequence 28468, A
32	146	9.3	698	13	US-10-424-599-116408	Sequence 116408, A
33	142.2	9.1	665	13	US-10-424-599-63982	Sequence 63982, A
34	132.2	8.5	345	13	US-10-424-599-95948	Sequence 95948, A
35	131	8.4	507	13	US-10-424-599-150803	Sequence 150803, A
36	126.2	8.1	1088	13	US-10-424-599-55315	Sequence 55315, A
37	121.4	7.8	1234	13	US-10-425-114-35022	Sequence 35022, A
38	115.4	7.4	748	13	US-10-425-114-3005	Sequence 3005, Appl
39	114.2	7.3	1491	13	US-10-424-599-107723	Sequence 107723, A
40	105.4	6.7	2147	13	US-10-424-599-76182	Sequence 76182, A
41	95.8	6.1	631	16	US-10-341-961A-234	Sequence 234, Appl
42	95.4	6.1	1904	13	US-10-425-114-22465	Sequence 22465, A
43	95	6.1	1793	13	US-10-424-599-47449	Sequence 47449, A
44	93	6.0	436	13	US-09-970-423-23	Sequence 23, Appl
45	89	5.7	1271	13	US-10-424-599-87597	Sequence 87597, A

#### ALIGNMENTS

RESULT 1  
US-09-992-901-1  
Sequence 1, Application US/0992901  
Patent No. US20020073446A1  
GENERAL INFORMATION:  
APPLICANT: Neff, Michael M.  
APPLICANT: Chery, Jeanne  
TITLE OF INVENTION: GENETICALLY MODIFIED PLANTS HAVING  
FILE REFERENCE: MODULATED BRASSINOSTEROID SIGNALING  
CURRENT FILING DATE: 2001-11-14  
PRIOR FILING DATE: 1999-12-14  
PRIOR APPLICATION NUMBER: US 60/124570  
PRIOR FILING DATE: 2000-03-16  
PRIOR APPLICATION NUMBER: US 60/124570  
PRIOR FILING DATE: 1999-03-16  
PRIOR APPLICATION NUMBER: US 60/170,931  
PRIOR FILING DATE: 1999-12-14  
PRIOR APPLICATION NUMBER: US 60/172,832  
PRIOR FILING DATE: 1999-12-20  
NUMBER OF SEQ ID NOS: 16  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 1: 1563  
LENGTH: 1563  
TYPE: DNA  
ORGANISM: Arabidopsis thaliana  
FEATURE:  
OTHER INFORMATION: Oligonucleotide  
US-09-992-901-1  
Query Match 100.0%; Score 1563; DB 9; Length 1563;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1563; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
DB 1 ATGAGAGAAGAAGAATGACAGCTGTCATTCAGAGGTTCTGTCTGTCTGAATCTTA 60  
1 ATGAGAGAAGAAGAATGACAGCTGTCATTCAGAGGTTCTGTCTGTCTGAATCTTA 60







QY 1052 GTGATGAGGCTCTCAGGCTCTGCGCTCAGCTGATGTCCTTACCAAGAGCATGCTTAA 1111  
 DB 1088 GCAAGAGATCTCTCAGCTCTGATGAGCAATCGACTTCTCTGTCGACAACTTTAATG 1147  
 QY 1112 AGCTTAAACGTTAGATGATCTTGAAGAGCTCTTAAAGTTGATCCACCAATGATG 1171  
 DB 1148 ACCTTAAGATTGTACCAATGATTAATTAAGAAACCTTCGCTCATCTCCAGAGATA 1207  
 QY 1172 CTACGATTCGACGCTTAATCGAGTGTGAAGCTAGAGAGGTACAAATCCCATGTGGCA 1231  
 DB 1208 TGTATGAGGCAAGCTCTTAAGATGTGATGCTTGAAGCATTAATTTCTGCTAAGA 1267  
 QY 1232 CGGAGCTTCTAATCCCAATCATACGCTCATCATGACCAAGCCATTTGGGGTAAATGAC 1291  
 DB 1268 CTCAACTCTTTTGGGATGAGCTGCTGTCATCATGACAGAAATCTGGGGAGAGATT 1327  
 QY 1292 TGAACGATTCATTCAGAGCTCGGTTGGGATGAGATGCGCGCTGCTGCCAAACACCCG 1351  
 DB 1328 ACCTTAACCTTCAATCCGATGAGATTAGCCCAAGCT-----CGCCCAAGCG 1375  
 QY 1352 TTGAGCTTCATACCGTTTGGCTCGAGTTCTGATCATGATGATGATGATCTTGTATAC 1411  
 DB 1376 TCGATTTCTTCTTCCCTTCGATTTAGGCTTCGATCTGTGTGGGCAAAATTTGGATGG 1435  
 QY 1412 TTGAGGCAAAATGACATCTGCTGTATGATCCAAAGCTTTCACCTTTCACTTGGCTCTA 1471  
 DB 1436 TTGAGGCAAAATGATGCTTCTTGTGATCATTCAAAGTTTGTGCTTATCTCAA 1495  
 QY 1472 CTATGAGCATGACCTTACCGCTCTGATGCTTATCTTATCTCAAGATGATGATCA 1531  
 DB 1496 ACRHATGATGCTCTCTAATCTATTTGTGATCTTGGAGCTCAATATGATGATCA 1555  
 QY 1532 CCTTCCGGA 1540  
 DB 1556 TTTTCAGAA 1564

RESULT 4  
 US-10-424-599-56717  
 ; Sequence 56717, Application US/10424599  
 ; Publication No. US20040031072A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: La Rosa Thomas J  
 ; APPLICANT: Kovalic David K  
 ; APPLICANT: Zhou Yihua  
 ; APPLICANT: Cao Yongwei  
 ; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
 ; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
 ; FILE REFERENCE: 38-21(53223)B  
 ; CURRENT APPLICATION NUMBER: US/10/424,599  
 ; CURRENT FILING DATE: 2003-04-28  
 ; NUMBER OF SEQ. ID NOS: 285684  
 ; SEQ. ID NO 56717  
 ; LENGTH: 1844  
 ; TYPE: DNA  
 ; ORGANISM: Glycine max  
 ; FEATURE:  
 ; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_22226C.1  
 US-10-424-599-56717

Query Match 18.6%; Score 291.4; DB 13; Length 1844;  
 Best Local Similarity 51.4%; Pled. No. 2.8e-79;  
 Matches 763; Conservative 0; Mismatches 696; Indels 24; Gaps 3;  
 DB 83 TGTCTCTGTTATGATGAGCAAGAAAGATTGAGACATTTCTTAAACAGGAATTC 142  
 DB 263 TGAATGGCTATGATGAGCAAGAAAGCTAGAGGCTCTTAAGAGGCAAGAGCTTTC 322  
 QY 143 GAGTCTCTCTTATCTTCTTATGAGAAATGTTAAAGATCTTGTGAATGATGCTTA 202  
 DB 323 AAGGAAACCCCTTACAGGATATGTGTGGGAGCTTAAAGAGATGTGAAATTTGCAATG 382

QY 203 AAGCTTCTTCTCATCTATGCGCTTCTCTCAAAATATCTTCTTACAGCTTCTCTTTT 262  
 DB 383 AAGCAAGATCCAAACCAATGATCTCTCCATGATGTGTCACAGTGTGTTGACATC 442  
 QY 263 ACCATCACTGAGAAATCTACGCTGCTATCTTCTGTTGGTGGTCCAACTTCC 322  
 DB 443 TGCATCGAGTCTCTCAACATGCAAGAAATCTTATTTGTTGGTTGACCAACCA 502  
 QY 323 GGTAAACGCTAGCCCATCTGATTTGATCAGAGATCTTCTTACGTGAGTTTACG 382  
 DB 503 GGGTACCTCCACAGATCTGAGCTGATCAAAAGTATCTTAACAGATCAGTATTTCC 562  
 QY 383 AGAAGATGAGACTCACCTTTGTTAAACAATTGAAGGAGATGACTACTGCTCA 442  
 DB 563 GAAAGCTGAAACCAATCACCTTGCCTTAAATTTACT--AGTACTGCTTGTAACTATG 619  
 QY 443 AAGTGAATAATGAGCTCATCATCGAAATATCATTTAGCCCTACTTTTCAATGAGATC 502  
 DB 620 ATGAGAAATGCAACAGACAGAAATTTAATTCCTGATTCAGTTTGAATAAT 679  
 QY 503 TTAAGTCTGTACAGCTGTGTGTTAAAGATGATGATGATGATGATGATGATGATG 562  
 DB 680 TGAAGATTAATGTTACCAATATCTTCAAAAGTTCATGATCTAATTAATCAAGTGAAG 739  
 QY 563 ATGATTAATCAGAAACGTTGAGTGAAGTGTATGATGATGATGATGATGATGATGATG 622  
 DB 740 GATGTTGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 799  
 QY 623 CTGAAGATTTATGATGAAACAGCTTTTGAAGTATGATGATGATGATGATGATGATG 682  
 DB 800 CTAGATGATTTATGCAAGAAACAGCTTTGGAAGTATGATGATGATGATGATGATGATG 859  
 QY 683 TTGATCTCAAGCTCAACAAATGCTTCTTGTGCTGAAGCTTTCAAAAGCTTCTTATTC 742  
 DB 860 TTCACTTCAAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 919  
 QY 743 CTGCTATGATTTTTCGACAAAGGAAATTTGAATCTCGAAGTTTACAGAGAGA 802  
 DB 920 CTGATGAGATTTTATCTTACTGACCAATCGAGATGAGAAATTTGATGATGATGATG 979  
 QY 803 TAAAGATTC-----GTTGTTAGCTGATGAGAGGCGGAGCAAAACGCTATG 853  
 DB 980 TAAAGCTCACTTACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1039  
 QY 854 ATGAGAAAGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 913  
 DB 1040 CTACTAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1099  
 QY 914 AAGCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 961  
 DB 1100 AAGAAACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1159  
 QY 962 TTTTCTTCCCGGAAACAGCACTTCTTATGCTGATGATGATGATGATGATGATGATGATG 1021  
 DB 1160 TCTACTTTCAGAGGAGAGACCACTTCACTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 1219  
 QY 1022 CCATGACCCGAGTGGAGGCAAGCAAGATGATGATGATGATGATGATGATGATGATGATG 1081  
 DB 1220 GTAGTACCTTATGAGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1279  
 QY 1082 GTGATGCTCTTACCAAGACATGCTGTTAAGCTTAAACGTTAGATGATGATGATGATGATG 1141  
 DB 1280 AAGCAACAAATTTGATGAGGCTTAAGTCACTTAAATGTCACATGATTTTGTGAAG 1339  
 QY 1142 AGCTTTAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1201  
 DB 1340 TGGTCTTAAAGCTTAAACCAAGCAAGTGTGCTTAAATGAAATGTTGACGAATATGA 1339  
 QY 1202 AGTGAAGAGGTAAACATCCCATGTGACAGAGCTTCTAATCCCAATCAATAGAGTTC 1261  
 DB 1400 AACTGGAAGACTATCACTTACTGCTGAGAGCAAGTTTCTTACCAACATGATGATGATG 1459  
 QY 1262 ATCATGACCAACCATTTGGGATATGATGATGATGATGATGATGATGATGATGATGATGATG 1321

Db 1460 CCCATGACCGCGAATCTTGGGGTATGATGTAATGTAATCAACCTGAGAGATTTCTG 1519  
 Qy 1322 ATGAGTGCCTGGGCTGCGCAACACCCCGTTGGCTTCAACCGTTGGCCCTGGAGTTC 1381  
 Db 1520 AAGAGTCTGTAAGGCCCAATGAGCAAGTTTCAATTTTTCATTTGGATGGGCTCTA 1579  
 Qy 1382 GTACATGATTTGGTGAAGATCTTGTACTATTCAGGCCAATTAACAAGCTGTATGA 1441  
 Db 1580 GAATATGATTTGGAACAAATCTTTCTCTTTGGAGGCAAGATGCTTTGTCAACATTT 1539  
 Qy 1442 TCCAAAGCTTCACTTCACTTGGCTCTACTATTCAGATGACCTAACCGTCTTATGT 1501  
 Db 1640 TACAATTTCTCATTTGAATTTCTCCAGCTTATGCTCATGCTCTGTATTAAGTATTTA 1699  
 Qy 1502 TGGTTATCTTCAACATGCTGACCAATCACCCTCCGAGATT 1544  
 Db 1700 CTCTTCAAGCCCAATATGCTGCTCATGCTATTTAAGTAAAGT 1742

RESULT 5

US-10-425-114-30156  
 ; Sequence 30156, Application US/10425114  
 ; Publication No. US20040034888A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Liu, Jindong  
 ; APPLICANT: Zhou, Yihua  
 ; APPLICANT: Kovalic, David K.  
 ; APPLICANT: Screen, Steven E  
 ; APPLICANT: Tabaka, Jack E  
 ; APPLICANT: Cao, Yongwei  
 ; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
 ; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
 ; FILE REFERENCE: 38-21(53313) B  
 ; CURRENT APPLICATION NUMBER: US/10/425,114  
 ; CURRENT FILING DATE: 2003-04-28  
 ; NUMBER OF SEQ ID NOS: 73128  
 ; SEQ ID NO 30156  
 ; LENGTH: 1886  
 ; TYPE: DNA  
 ; ORGANISM: *Oryza sativa*  
 ; FEATURE:  
 ; OTHER INFORMATION: Clone ID: UC-OSFLCVP143F07\_FLI  
 US-10-425-114-30156

Query Match 17.6%; Score 274.8; DB 13; Length 1886;  
 Best Local Similarity 51.2%; Pred. No. 4, 4e-74;  
 Matches 751; Conservative 0; Mismatches 697; Indels 18; Gaps 4;

Qy 94 TGGTGGAGACCAAGAAAGTGAAGACATTTCTCTAAACAAGAAATTCGAGGCTTCCT 153  
 Db 225 TGGTGGAGACCGCGCGCTGAGCGCGCTCCGCGCGAGGCAATCCGCGCAACCG 264  
 Qy 154 TATCATTTCTTCACTCGGAATGTTAAAGACTTGTGAAATGATGCTTAAAGCTTTCT 213  
 Db 285 TACCGCTCTTCAACCGCGAGCGTCCGGAAGAGTCCGCTCAACCGGAAGCCGGAAG 344  
 Qy 214 CATCTATGCTTTT---CTCTCAATATTTCTCTAGAGTCTCTTTTTCATCATAC 270  
 Db 345 AAGCCGCTGCGCTGCGCTGCAACGATCATCCGCGGCTGCTCCGATGTTACGAAA 404  
 Qy 271 TGAAGAAAATCTACAGTGTCAATTTGTTGGTTGGTTCGATCAACTTCCGTTAAG 330  
 Db 405 GCGGTTGAGAGAGACCGGAACATCATTTCACTGGTTTGGCCCAACCGAAGATGATG 464  
 Qy 331 GTAGCCGATCTGATTTGATCAGAGATCTTCTCTAAGTCTGAATTTCTAAGAAAGAT 390  
 Db 465 ATTTCAAGACCTGATCAATTAAGGAAGTATGTCTAATTAAGTTTGGCCACTAAGCAAA 524  
 Qy 391 GAAGCTCAACCTTTGGTTAAACAATTGAAGGAGATGACTACTAGTCTCAAAAGGTAA 450  
 Db 525 CAAAGCTTACCGGCTCGGAAGATGCTAGCTCCGGAAGTGTAAAGCAAGAGCGAG 584

Qy 451 AAATGGGCTCATCATGAAAAATCATTTAGCCCTACTTTTCATATGAGAAATCTTAAAGTTG 510  
 Db 585 AAATGGGCAAGACCGGAGAAATCTGAAATCTTCCCTTTTACACGAGAAATTAAGCGG 644  
 Qy 511 CTGTGACAGATTTGTTGAAGATGATGATGATGATGATGATGATGATGATGATGATG 570  
 Db 645 ATGCTGCAATTTTCTTACTGCTGACCGAAATGTTTCAAGATGGGAAATTTCAATG 704  
 Qy 571 TCAGAAAACGCTGAAGTGAAGTATGATGATGATGATGATGATGATGATGATGATGATG 630  
 Db 705 TCTATGAGAAAGATGACAGATGATGATGATGATGATGATGATGATGATGATGATGATG 764  
 Qy 631 GTTATTAATGAACGCTTTTGAAGTATGATGATGATGATGATGATGATGATGATGATG 690  
 Db 765 GTCATATCAAGACAGATTCGTGAGACGATGATGATGATGATGATGATGATGATGATG 824  
 Qy 691 CAAGCTCAACAAATGCTTTTGTGCTGAAGCTTTTCAAAAAGCTTCTCTGCTAT 750  
 Db 825 CAAGCAGATGACCGCAACGATTAATACAGCTTTTGGACATTTTATACAGATAT 884  
 Qy 751 AGATTTTTCGACAAAGAGAAATTTGAAGTCTCGAAGTTAGCAAGAGATTAAGAA- 809  
 Db 885 TGGTTCTTACCAATTAACAAACAGAAAGTTGAGAAATTAAGAAAGAGGTCACGAAA 944  
 Qy 810 -----GTGCTTGAAGCTGATGAGCGGCGGAGAACAAACGCTATGATGAGAA 861  
 Db 945 CTCTACAGAGAAATTAATGAGAAAGAGAGGCGCTTTAATAATGATGAAACCGATAT 1004  
 Qy 862 GGGGAAGATGTAAGAGACCGCGCGAGAGATTTGTTGGATTAATGATTCAGGCAAG 921  
 Db 1005 GGTGACTTGTGGGCTTATGTTGTAAGTCAAAATGAGGAGTCAATGAGGAAGCGAAA 1064  
 Qy 922 AATG---TGACGTTTCAAGAAATTTGTAAGAGATGTAAGCTTTTCTCGCGGAGAA 978  
 Db 1065 CTAGAAATGACTACGAGAAATTAATGAGAAAGCAAGCTATTTTATTTGACGAAATG 1124  
 Qy 979 CAGACAACTTCTAATCTGCTGACGAGACGACATCTTGTATCCATGACACCGAGTGG 1038  
 Db 1125 GAGCAATCATGATTTGCTCATTTGACATTAATTTGCTAAGTATGACCCAGAGTGG 1184  
 Qy 1039 CAGCCAAAGCATGATGAGTCTCTAGGCTCTGCGGCTCACTGATGTCCTTACCAAG 1098  
 Db 1185 CAGAGCGAGCAGAGAAAGAGTGTACACCACTTTGG--AAAGAACACACGAGCTAT 1241  
 Qy 1099 GACATGCTGTTAAGCTTAAACGTTGAGTATGATCTTTGAACGAGTCTTTAAGTTGAT 1158  
 Db 1242 GATAGCTTAAGCTGCTGAAGATTTGAATTTGATGATGATGATGATGATGATGATGAT 1301  
 Qy 1159 CCACCAATTAAGTACGATTCGACGCGCTAAATGAGATGTAAGCTGAGAGGTACAAA 1218  
 Db 1302 CCGCCAGATGTTGTTTCTTACACGAAATTAAGCAAGAAATGAGGCTCGGCGGATPAAA 1361  
 Qy 1219 ATCCCATGAGCAGAGATTCATATCCCAATCAATAGCGGCTATGACCAAGCATT 1278  
 Db 1362 TATCCGCTGAAGTACCTTATGTTGCCATTTTATTAACAGATGATCCGATAT 1421  
 Qy 1279 TGGGTAATGAGATGAGCAATTAATCAATCAATGATGATGATGATGATGATGATGATG 1338  
 Db 1422 TGGGAAAAGATGACAGTAAATTAATCAATCAGGAGATTTGCTGATGATGATGATG 1481  
 Qy 1339 GCCAAACACCCGTTGCTTCAATGATGATGATGATGATGATGATGATGATGATGATG 1398  
 Db 1482 ACGAATATCAACCTCTTCTTCTTCAATTTGATGATGATGATGATGATGATGATGATG 1541  
 Qy 1399 AATCTGCTATCTTCAAGCCAAATTAATCAATGATGATGATGATGATGATGATGATGAT 1458  
 Db 1542 AACTTTGACATTAATGAGACCAAGATGATGATGATGATGATGATGATGATGATGATG 1601  
 Qy 1459 CACTTGGCTCTTATTAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1518  
 Db 1602 GAGCTTCAACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1661  
 Qy 1519 GGTGACCAATCACTTCCGAGATT 1544

Mon May 24 08:10:08 2004

us-09-992-901-1.rmpb

Page 6

Db 1662 GGTGCAAAATTTAGCTGAAGAAAT 1667

RESULT 6

US-10-425-114-22323  
; Sequence 22323, Application US/10425114  
; Publication No. US20040034888A1  
; GENERAL INFORMATION:  
; APPLICANT: Liu, Jindong  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Screen, Steven E  
; APPLICANT: Tabaska, Jack E  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(5313)B  
; CURRENT APPLICATION NUMBER: US/10/425, 114  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 73128  
; SEQ ID NO: 22323  
; LENGTH: 1836  
; TYPE: DNA  
; ORGANISM: Oryza sativa japonbare  
; FEATURE:  
; OTHER INFORMATION: Clone ID: LIB3431-054-B3\_FLI  
US-10-425-114-22323

Query Match 17.5%; Score 273.2; DB 13; Length 1836;  
Best Local Similarity 51.2%; Pred. No. 1,4e-73;  
Matches 750; Conservative 0; Mismatches 698; Indels 18; Gaps 4;

QY 94 TGGTGAAGACCAAGAAATGAGAAATCTTCTTAAACAGAAATGAGTCTCT 153  
DB TGGTGAAGACCGGCGGCTGAGAGCGCCCTCCGGGCGAGGCAATCCGGGCAACCG 239  
QY 154 TATCATTTCTTCAATCGAAATGTTAAAGAACTTGTGGAATGATGCTTAAAGCTTCT 213  
DB TACCCGCTTCAACCGGCGAGTCCGAGAGAGTCCGGGCTCAACCGGAGCGCGGAA 299  
QY 214 CATCTATAGCTTT---CTCTCACAATATCTTCCAGAGTCTCTCTTTTACATGAC 270  
DB AAGCGCTGCGGCTGCGTCCGACGACATCATCCGCGCTGCTCCGATGTTCCAGCAA 359  
QY 271 TGGAGAAAATCTAGAGTCTCAATTTCTGTTTGTTCGTCGCACTTTCCGTTAA 330  
DB GCCGTTGAGAGACAGGGAACCATCTTCACTTGTGTTGCGCCACGCGCAAGTGA 419  
QY 331 GAGCGCATCTGATTTGATCAGAGAGATCTCTCTAAGTCTGAGTTTCAAGAGAA 390  
DB ATTGAGACCTTGAATCAATTAAGGAAATTAATGCTATAGTTTGGCACAATAGCAA 479  
QY 391 GAAGCTCACTTTGGTTAAACATTGAAGGCGATGAGACTTAACTTGAAGTGA 450  
DB CCAAGCTTACCCGCTCTGGAAGTGTCTAGCTCCGAGTTGTAAGCTATGAAGCGAG 539  
QY 451 AATGAGCTCATCTGAAATCATTAAGCCTTACTTTTCAATGAGAGATTTAGTTG 510  
DB AATGAGCAAGACCGGAGAAATCTGAATCTGCTTTCACAGAGAAATTAAGCGG 599  
QY 511 CTGTGACCAATTTGTGTTGAAGTGTGAATGATGATGATGATGATGATGATGAT 570  
DB ATGTGACCAATTTTCTTCACTGTCAGCAAGAAATGTTTCAAGATGAGAGATTTCA 659  
QY 571 TCAGAAACGCTGAAGTGTGAGTATGATGAGAGTGTGAGTGTGAGTGTGAGTGT 630  
DB TCTATGAGAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 719  
QY 631 GTTATTTAGTGAACAGCTTTTGAAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 690  
DB GTCATATCAAGACAGCAATTTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 779

QY 691 CAAGCTCAACAAATGCTTCTTTGCTGAAGCTTTTCAAAAAGCTTATCTCTGCTAT 750  
DB CAGAGAGATGCGCCGAAACGATTAATACAGCTTTTGGACATTTTATACAGATAT 839  
QY 751 AGATTTTTCGACAAAGGAAATTTGAATCTTCCGAAATTTAGACAGAGATTAAGAA 809  
DB TGGTCTTCAACCACTTAAACAAAGAGGTTGAGAAATTTGAAAGAGAGTCAACAAA 899  
QY 810 -----CTGCTTTGAAGCTATGAGAGCGGAGACAAAACGCTATAGATGAGAA 861  
DB CTCTACAGAGAAATATTTGAAGAGAGAGCGGCTTATTAATATGTTGAACCAATAT 959  
QY 900 CTCTACAGAGAAATATTTGAAGAGAGAGCGGCTTATTAATATGTTGAACCAATAT 959  
DB GGTGACTTGTGGCTTATTTGTGAGAGTCAAAATTAAGGAGTCAAAATGGAAGCAGAA 1019  
QY 922 AATG---TGAGGCTTCAAGACATTTGAGAGAGTAAAGCTTTTCTTCCGCGGAAA 978  
DB CTAGAGATGATCTAGAGAGAAATTTTGAAGATGCAAGCTATTTTATTTTGAAGAA 1079  
QY 979 CAGACACTTCTAATCTGCTGAGTGAAGACACATCTTCTATTCATGACCCGAGTGG 1038  
DB GAGACAACTCAGTATTTGCTCATTTGATGATGATGATGATGATGATGATGATGAT 1139  
QY 1039 CAGGCAAGACAGTGAAGAGTCTCAGAGTCTCAGAGTATGCTTACCAAG 1098  
DB CAGAGAGAGACAGAGAGAGAGTCTTACACCACTTTGG---AGAAACACACAGACTAT 1196  
QY 1099 GACCATGCTGTTAAGCTTAAACGTTGAGTATGATCTTGAACAGAGCTTTAAGTTGAT 1158  
DB GATAGCTTAAAGCTGCTGATGATGATGATGATGATGATGATGATGATGATGAT 1256  
QY 1159 CCACCAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1218  
DB CCGCAGTGTGTTTCTTACAGAGAGAGTCTTACAGAGAGTCTTACAGAGAGTCTT 1316  
QY 1219 ATCCATGTGACAGGAGCTTCTAATCCATCAATGAGTCTTCAATGACAGCAAT 1278  
DB TATCCGCTGAGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1376  
QY 1279 TGGGATATGAGTGAAGATGATGATGATGATGATGATGATGATGATGATGATGAT 1338  
DB TGGGAGAGAGATGAGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 1436  
QY 1339 GCCAAACACCCGTTGCTTCAATACGTTTGGCTCGAGTGTGATGATGATGATGAT 1398  
DB ACGAAGTATCAGACCTCTTCTTCCATTTGATGAGTGGGCTCCCGAATCTGCAATCGGCG 1496  
QY 1399 AATCTGCTATCTCTCAGGCAATGATGATGATGATGATGATGATGATGATGATGAT 1458  
DB AATCTGCTATCTCTCAGGCAATGATGATGATGATGATGATGATGATGATGATGATGAT 1556  
QY 1459 CACTTGTCTCTTCAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1518  
DB GAGCTTCAACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1616  
QY 1519 GGTGACCAATCACTTCCGAGATT 1544  
DB GGTGACCAATTTAAGCTGAAGAAAT 1642

RESULT 7

US-10-425-114-13396  
; Sequence 13396, Application US/10425114  
; Publication No. US20040034888A1  
; GENERAL INFORMATION:  
; APPLICANT: Liu, Jindong  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Screen, Steven E  
; APPLICANT: Tabaska, Jack E  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With

TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
 FILE REFERENCE: 38-21(5313)B  
 CURRENT APPLICATION NUMBER: US/10/425,114  
 CURRENT FILING DATE: 2003-04-28  
 NUMBER OF SEQ ID NOS: 73128  
 SEQ ID NO: 13396  
 LENGTH: 1877  
 TYPE: DNA  
 ORGANISM: Oryza sativa japonbare  
 FEATURES:  
 OTHER INFORMATION: Clone ID: JC-OSLELIB3474019601\_FLI  
 US-10-425-114-13396

Query Match 17.5%; Score 273.2; DB 13; Length 1877;  
 Best Local Similarity 51.2%; Pred. No. 1.4e-73;  
 Matches 750; Conservative 0; Mismatches 658; Indels 18; Gaps 4;

94 TGGTGAAGACCAAGAAATGGAACATTTCTTAAACAAAGGAATTCGAGTCCCT 153  
 219 TGGTGAAGACCCGCGCGCTGAGGCGCCCTCCGCGCGAGGCAATCCGCGCAACCG 278  
 154 TATCATTTCTTCATCGAAATGTTAAAGACTTGTGATGATGCTTAAAGCTTCTCT 213  
 279 TACCGCTCTTACCGGCGAGCGTCCGGAAGCGTCCGCTCAACCGGAGCGCCGGAAG 338  
 214 CATCTAATGCTTT---CTCTCACAATATCTTCTTCAAGTTCCTTTTTCACATCAC 270  
 339 AAGCGCTGCGCGCTGCGCTGCGACGACATCAATCCGCGCGCTGCGCGAGTTCAGCAA 398  
 271 TGGAGAAAATCTAGCGTCTACATTTCTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 330  
 399 GCGGTGAGGAGCAAGGGAACCATCATCTGTTTGGCCCAAGCGCAAGGATGATG 458  
 331 GTAGCCGATCTGATTTGATCAGAGATCTTCTCTAAGTGTGATGCTTCAAGAAAT 390  
 459 ATTTCAGACCTCGAATCAATAGGAGATGTTATGTTCTAATTAAGTTGCGCACTATG 518  
 391 GAAGCTCAACCTTGTGTTAAACATCTGAAGCGATGAGTACTTAACTTCAAGTGA 450  
 519 CCAAGACCTACCCGCTCGGAAGTGTCTAGCTCCGAGTGTATAGTATGAAGCGAG 578  
 451 AATGCGCTCATATCGAAATATCATAGCCCTACTTTTCAATGAGATCTTAAATG 510  
 579 AATGCGCAAGACACCGGAGATCTGAACTCTGCTTTTCAACGAGAAATTAACCG 638  
 511 CTGTACCACTGTGTGTAAGAGTGAATGATGATGATGATGATGATGATGATGATGAT 570  
 639 ATGCTGCCAGTTTCTTCTACTGCTGACGGAATGTTTCAAGATGAGGAATTTCAATG 638  
 571 TCAGAAACGATGAAATGAGTGAATGCTATGAGTGTGATGATGATGATGATGATGAT 630  
 699 TGTATGAGAGATGCAAGATGATGATGATGATGATGATGATGATGATGATGATGAT 758  
 631 GTATATAGTAAAGCGCTTTGAGAGTGAATGAGTGAATGAGTGAATGAGTGAATGAG 690  
 759 GTATATCAAGACGATTCGATGAGTGAATGAGTGAATGAGTGAATGAGTGAATGAG 818  
 691 CAAGCTCAACAAATGCTTCTTGTGCTGAAGCTTTTCAAAAGTCTTCAATCTGAGTAT 750  
 819 CAGGAGAGTGAAGCGCAAGCAATATACAGCTTTCGAGCAATTTTATACAGGATAT 878  
 751 AGATTTTTCGCAAGAGGAAATTTGAAGTCTCGAAGTGAATGAGTGAATGAGTGAAG 809  
 879 TGGTCTTACCAACTTAAACAAACAGAGGTTGAAGAAATTTGAAGAGAGTCAAGAA 938  
 810 -----GTGCTGTTGAAGCTGATAGAGCGCGGAGCAAAACGCTATAGTGAAGAA 861  
 939 CTCTTACGAGAAATATTTGAAAGAGAGCGGCTATTTAAATATGAGTGAAGTGAAT 998  
 862 GGGGAAAGATGTAAGAGCGCGCGGAGAAATTTGTTGAGTATTAATTTACGAGCAAG 921  
 999 GGTGACTTGTGGGCTTATTTGTTGAGTCAATATGAGGAGTCAATATGAGGAAAGCA 1058

922 AATG---TGACGTTTCAAGCATTTGAGAGAGTGTAAAGCTTTTCTTCCGCGGAAA 978  
 1059 CTAGAAATGACTACGAGCAAGAAATTAAGAAATGCAAGTATTTATTTTGAAGAAAG 1118  
 979 CAGCAACTTCTAATCTGCTGAGTGAACGACATCTTGTCTATCCATGACCCGAGTGG 1038  
 1119 GAGACATCATGATGATGCTGCTCACTGAGCAATTAATGTTGTTAGTATGACCAAGTGG 1178  
 1039 CAGCCAAAGCACTGATGAGGCTCTGAGGCTTCCGCGCTCAGTATGCTTACCAAG 1098  
 1179 CAGGCGAGGCAAGAGAAAGAGTCTACACACTTTGG---AAGAACACACACAGCTAT 1235  
 1099 GACCATGCTTAAAGCTTAAAGCTTATGATGATGATGATGATGATGATGATGATGAT 1158  
 1236 GATGACTTAAGTCTGCTGAGAAATGTTAAATATGATGATGATGATGATGATGATGAT 1295  
 1159 CCACCAATAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1218  
 1296 CCGCAGAGTGTGTTCTTGACCAAGCAATACAGGAAATGAGGCTCGCGGCAATCAA 1355  
 1219 ATCCAGTGGCAGAGGCTTCTATCCCAATCAATAGCGGTCCATCATGACCAAGCAT 1278  
 1356 TATCCGCTGAGTGAAGTACCTTATGTTGCTTATTTATTTATTCACATGATCCGATAT 1415  
 1279 TGGGTAATGACGTGAAGCAATTCATCAAGCTCGGTTTGGAGTGAAGTGGCGGTGCT 1338  
 1416 TGGGAAAGATGAGAGGATGATTCATTCAGGAGAGGTTTCTGATGATCTTCAACGCA 1475  
 1339 GCAAAACCCCGTGTGCTTCAATCGCTTGTGCTCGGATGTTGATCATGATGATGAT 1398  
 1476 ACGAAGTATCAGACCTCTTCTTTCATTTGATGAGTGGGTTCCGAAATCTCATCGG 1535  
 1399 AATCTGTATCTTACAGGCAATGACACTGCTGTAATGATTCACAGCTTACCTTT 1458  
 1536 AACTTGTGATGATGAGAGCAAGTGTATCTGTAATCTTCAAGGCTTCTCTTT 1595  
 1459 CACTTGTGCTTACTTATCAGCATGACCTACCTGCTTATGTTGCTTATCTTCAACAT 1518  
 1536 GAGCTTACCATGATGATCAAGCAATTCATCTGATTAATCTTCAACCAAGCAT 1655  
 1519 GGTGACCAATCACTTCCGAGAT 1544  
 1656 GGTGCAAAATTAAGCTGAAGAAAT 1681

RESULT 8  
 US-10-425-114-1043  
 Sequence 1043; Application US/10425114  
 Publication No. US20040034888A1  
 GENERAL INFORMATION:  
 APPLICANT: Liu, Jindong  
 APPLICANT: Zhou, Yihua  
 APPLICANT: Kovalic, David K.  
 APPLICANT: Screen, Steven E  
 APPLICANT: Tabaska, Jack E  
 APPLICANT: Cao, Yongwei  
 TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
 FILE REFERENCE: 38-21(5313)B  
 CURRENT APPLICATION NUMBER: US/10/425,114  
 CURRENT FILING DATE: 2003-04-28  
 NUMBER OF SEQ ID NOS: 73128  
 SEQ ID NO: 1043  
 LENGTH: 1853  
 TYPE: DNA  
 ORGANISM: Zea mays  
 FEATURES:  
 OTHER INFORMATION: Clone ID: 700100628\_FLI  
 US-10-425-114-1043  
 Query Match 17.1%; Score 267.8; DB 13; Length 1853;  
 Best Local Similarity 51.4%; Pred. No. 6.6e-72;  
 Matches 679; Conservative 0; Mismatches 627; Indels 15; Gaps 2;

QY 228 CTCTCAAAATTTCTTCTAGAGTCTCTTTTTCATCTCACTAGTGAAGAAAATCTAGCG 287  
 Db 310 CTGCCACGACATCACCCCACGCTGACGCCATGATCAACAGCAATCAAGAAATACGG 369  
 QY 288 TGGTAACTTTCTGGTTGGTTGGTTCACCTTCCGGTTAACGGTAACCGATCTGATT 347  
 Db 370 GAAACTATCGTTCACTGGTTCGGCCCAACCAAGGGTGAATGATCTCGAACCGAGTT 429  
 QY 348 GATGAGAGATCTCTCTAAGCTCTGAGTTCTACGAGAAATGAAGTCACTACCTTTGT 407  
 Db 430 AGTCAAAAGAGTGTCTTAATAGTTGGCCACTTTGGCAACCAAGAGTAAGCCGAT 489  
 QY 408 TAAACAATTGAAGCGAGTACTTACTAGTCTCAAAAGTGAAGAAAATGGCTCATCG 467  
 Db 490 TGGAGGTTCTAGCCCAACGGGCTTGAATCATGATGTGAAAATGGGCAACACAG 549  
 QY 468 AAAATCATAGCCCTCTTTTCAATAGAGAACTTAAGTTGCTTACAGTTGTGT 527  
 Db 550 GAGATTTCTTAATCTGATTTCAATCAAGAAAATTAAGGGATGATGCCAGTATTTTC 609  
 QY 528 GAAAGTGTACTGATATGTTGATTAATGTTCCGATTAAGTTATCAGAAAACGTTAAGT 587  
 Db 610 TACTGCTGATTAATGATTAATGATTAATGATTAATGATTAATGATTAATGATTAATG 669  
 QY 588 TGAAGTATGATCTATGATGATGTTTCAAGTTTGAAGTGAAGTGAAGTGAAGTGAAG 647  
 Db 670 TGAATATGATCTGAGCTGATGATGATGATGATGATGATGATGATGATGATGATG 729  
 QY 648 TTTTGAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 707  
 Db 730 GTTTGGAGCACTATCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 789  
 QY 708 TCTTTGCTGAAGCTTTTCAAAAGTCTTCAATCTGCTGATTAATTTTTCGACAG 767  
 Db 790 ACGCTCATCAATCTGCTGATGATTAATTAATTAATTAATTAATTAATTAATTAATTA 849  
 QY 768 AGGAAATTTGAAGCTGGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 827  
 Db 850 AAACAAAGAGGATGAGAGCAATGATGAGAGTCCGAAATTTCTCCGAAATAT 909  
 QY 828 AGAGCGCGAGAGCAAAACGCTATGATGAGAGGAGGAGGAGGAGGAGGAGGAGGAG 879  
 Db 910 TGGAAAGAGAGAGGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 969  
 QY 880 ----CCGGCGCGAGAGATTTGTTGGATTAATGATTAATGAGGAGGAGGAGGAGGAG 935  
 Db 970 ATTAAGTGAAGTGAAGCAAGAGCAATTAATTAATTAATTAATTAATTAATTAATTA 1029  
 QY 936 GACATTTGAGAGAGTGAAGAAAGCTTTTCTGCGCGAGAAACAGAACTTTCTAATCT 995  
 Db 1030 AGATGATTAAGAGAGGAGGATTAATTTTCTTTGAGGATGAGAGCAATCAATCTCT 1089  
 QY 996 GCTGAGTGAAGCACTCTTCTATGATGAGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1055  
 Db 1090 GCTTACTTGAACCTTATTTGTTGATGAGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1149  
 QY 1086 TGAAGTCTCAGAGGCTGCGGCTCAGCTGATGTCCTTACCAAGAGCAATGCTTAACT 1115  
 Db 1150 AGAGTTTGAAGCACTTTGG---AAGAACACACAGATTAATGATGCTTGAAGCGCT 1206  
 QY 1116 TAAAGCTTGAATGATCTTGAAGCAAGTCTTTTAAAGTTGATTAATTAATTAATTA 1175  
 Db 1207 CAAGCTGTAACATGATTTCTAATGAGTCTTAAAGTTGATTAATTAATTAATTAAT 1266  
 QY 1176 GATTGAGAGCGCTAAATGAGATGAGAGCTGAGAGGAGTGAAGAAATCCCATGTCAG 1235  
 Db 1267 AACCAAGAACTTTAAGAGAAATGAGAGTGGAGTGGAGTGAAGTCTGAGAGAGTGA 1326  
 QY 1236 GCTTTAAATCCCAATCAATGAGGTCATGATGAGCAAGCAATTTGGGATTAATGATGA 1295  
 Db 1327 GCTCTTCTGCGCGCTCATCTTCAATCAATGATTCGAGCAATTTGGGAGAAAGCAAG 1386

QY 1296 CGAATTAATCCAGCTCGGTTTGGAGATGAGTCCCGGCTGCTCCAAACACCCCGTTGG 1355  
 Db 1387 CGAGTTCACCCAGAAAGGTTTGGCCAGAGGATCTCCAGCCCAACAGGATCAGGCTGC 1446  
 QY 1356 CTTCATACCGTTGGCTCGAGGTTCTGATGATGATGATGATGATGATGATGATGATGAT 1415  
 Db 1447 TTTCTTCCGTTCCGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1506  
 QY 1416 GGCCAATTAACATCTGCTGATTAATGATTAATGATTAATGATTAATGATTAATGAT 1475  
 Db 1507 AGCCAGATGAGAGGATGAGAGGATGAGAGGATGAGAGGATGAGAGGATGAGAGG 1566  
 QY 1476 TCAGATGAGAGGATGAGAGGATGAGAGGATGAGAGGATGAGAGGATGAGAGGAT 1535  
 Db 1567 CACCCAGGCGGCTTACCGCTGATTAATCACTGACCTCAGAGCGGTGCTAGATAGGT 1626  
 QY 1536 C 1536  
 Db 1627 C 1627

RESULT 9  
 US-10-424-599-85977  
 ; Sequence 85977, Application US/10424599  
 ; Publication No. US20040031072A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: La Rosa, Thomas J  
 ; APPLICANT: Kovalic David K  
 ; APPLICANT: Zhou Yihua  
 ; APPLICANT: Cao Yongwei  
 ; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
 ; FILE REFERENCE: 38-21(53223)B  
 ; CURRENT APPLICATION NUMBER: US/10/424,599  
 ; CURRENT FILING DATE: 2003-04-28  
 ; NUMBER OF SEQ. ID NOS: 285684  
 ; SEQ ID NO 85977  
 ; LENGTH: 1922  
 ; TYPE: DNA  
 ; ORGANISM: Glycine max  
 ; FEATURE:  
 ; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_48650C.1  
 ; US-10-424-599-85977

Query Match 16.3%; Score 254.2; DB 13; Length 1922;  
 Best Local Similarity 50.6%; Pred. No. 1.2e-67;  
 Matches 759; Conservative 0; Mismatches 718; Indels 24; Gaps 5;

QY 72 AGTGAAGGATGATCTCTGTTATGAGAGCAAGAAAGATTGAAGAACTTTCTTAA 131  
 Db 107 ACTGAATATCTTAATCTGCTGCTGCTTGGCCAAAGGTTGAAGGCTTTAGAGC 166  
 QY 132 ACAAGAAATGAGGCTCTCTTATCATTTCTTATGAGAAATGTTAAAGAACTTTGG 191  
 Db 167 GAAAGTTTCAATGAGAGGATCACTTCTCATCACTCACTCACTCACTCACTCACTTT 226  
 QY 192 AATGATCTTAAGCTTCTTCTCATCTGATGAGGCTTCTCTCAAAATTTCTTAAAGT 251  
 Db 227 GCAACAAACACACTCAATTTCCAACTTTTCTCTATGAGATGATGCTGCTGCT 286  
 QY 252 TCTCTCTTTTACCATCACTGAGAAATTAACGCTGCTAATTTCTGTTGGTGG 311  
 Db 287 ATCTCTGTTTACATCACTATGCAAAATGAGCAAAATGAGCAAAATCTTTTGGAGG 346  
 QY 312 TCAACTTTCGCTTAAAGTACCGATCTGATTTGATGAGAGATCTTCTTAAGTC 371  
 Db 347 TAGAACAACAAAGGATGATTAAGAGCCAAATCACTCAAGAGTCTTCAACAAT 406  
 QY 372 TGAATTTCAAGAGAGATGAGAGTCACTCTTTGTTAAACAATTGAAGGAGTGAAT 431  
 Db 407 TCAATGCTTCAAGAGCAAAAGTTAGTGAATGATGATGATGATGATGATGATGATGAT 463  
 QY 432 ACTTATGCTCAAGAGTGAAGAAATGGGCTCATGATGAGAAATCATTAAGCCCTACTTTCA 491

Db 464 GTTAATATAGAGGTGCAAGTGGGCTAAACATCGAAGATTATGACCACCATTTCA 523  
Qy 492 TATGAGAACTTTAAGTCTTGTCACAGTTGTGTGAAGAGTGTACTATATGTGTGA 551  
Db 524 CTCAGAAAAATTTGAAAAATATGCTACCGCATTTTCTCAAGCTGCCATATATGATTAG 583  
Qy 552 TAAATGTCCTGATTAATATCAAGAAACGTGAGTGTGAGTATGTATAGTGTGT 611  
Db 584 CATGTGGAAGGAAAGTGTGTCTATCAATGAGAAATGTGATGACATTTGGCTTCT 643  
Qy 612 TCAGATTTGACTGAAGATGTTATTTAGTGAACAGCTTTTGAAGTATGATGAAGTGG 671  
Db 644 TCAGATTTAACTCGTATGTTAATTTCTCAACAGCATTTGGAAGCAGCTATGCAAGAG 703  
Qy 672 TCGAGCAGTTTTCGACTCAAGCTCAACAAATGCTTTGTGCTGAGCTTTTCAAA- 730  
Db 704 AGAAAAATTTTTCGAAAATCGAGATGCAAGGCTATCTTTTAATGCGAGAAAGTCA 763  
Qy 731 -----AAGCTTCAATTCCTGCTATGATTTTTCGCAAGAGGGAATTTGAAGT 781  
Db 764 GAACATACCAATATTGCGGCATCTACGTACAACTACCAAGAGATGAGCAATGGA 823  
Qy 782 CTCGGAATTTAGACAGAGATAGGAAGTGTGTGAGCTGATGAGCGCGGAGA- 840  
Db 824 AAGAGATACGAGATTCATTTGAGGGTATCAAAAAAGAGAGAAACCATGAGAA 883  
Qy 841 -----CAAAAGCTATAGATGAGAGAGGGAAGATGTAAGAGCGCGCGAGATTT 896  
Db 884 TCGTGAACCTTATATGAAATTTATTAAGCATCTTTGGAATCAATACAAAGAAAT 943  
Qy 897 GTTGGGATT--AATGATTCAAGGCAAGAAATGTGACGGTTCAGACATTTGAGAGAGT 953  
Db 944 CAGGAGACATGGAACAGTAGGGCTGTGATGACCAAGCAAAAGTATTTGAGATG 1003  
Qy 954 TAAAGCTTTTCTTGCGCGGGAACAGCAACTTCTATCTGTCAGTGAAGCAACAT 1013  
Db 1004 CAGCTATTTTACCTGCGAGGCAAGACACCTTCTCTGCTGTGTGACATAGT 1063  
Qy 1014 CTGCTATTCATGCAACCGGAGTGGCAGGCAAAAGCAGTATAGGTCTCAGGGCTG 1073  
Db 1064 CTATATGGCTAGGTATCCAGATGCGAAGCAGAGCAAGGAGCAAGTTTTCAGTTT 1123  
Qy 1074 CGGCTCAGGTATGTCCTACCAAGACATGTGTTAAGCTTAAAGCTTGAATGAT 1133  
Db 1124 TGGAAACCAAAAT--CCAAACATTTGAGGGTTAAGTGAAGCTTAAATGTGACATGAT 1180  
Qy 1134 CTGGAACGATCTTTAAGTGTGATCCAAATGATGCTAAGTTCAGTGCAGCGTAAATC 1193  
Db 1181 TTTAATACAGGTTCTCAGGCTATATCTCCAAACAATTTCTTGTGCGGCACTCAAA 1240  
Qy 1194 GGAATGAGCTGAGAGGGTACAAATCCCATGTGGCAGAGCTTTCTATCCCAATCAT 1253  
Db 1241 GGATGTGAACCTTGAATACTATCTACCTGAGGAATACGATTTCCATGCCAATATT 1300  
Qy 1254 AGCGTCCATCATGACCAAGCATTGGGGTAAATGACGTGAACGAATTCATCCAGCTCG 1313  
Db 1301 ATTATTCATCATGATGATGATGATCTGGGGTGAATGCAAAAGAGTTCAACCTGAAG 1360  
Qy 1314 GTTTGGGAGTGAAGTGGCGGCTGCTGCCAAACCCCGTTGGCTTCATACCGTTGGCT 1373  
Db 1361 GTTTCTGAGGAATTTGCTAAGGCAAAAGGCCAAATTTGTTTATTCATTGTGATG 1420  
Qy 1374 CGAGTTGCTACATGATGTCAGATCTGTATATCTCAGGCCAAATTTGACACTCG 1433  
Db 1421 GGGTCTCTAGAAATATGATTTGGCCAAACCTTGCCTTAATGGAACAAAGATAGCTTATC 1480  
Qy 1434 TGTAAATGATCAACGCTTCACTTCACTTGGCTCTATCTTATAGATGACCTTACCT 1493  
Db 1481 ATTGCTTCTGCAACATTTCTCATTTGAGCTTTCGCGGTCTATGAAATGCTCCAAAGGT 1540  
Qy 1494 CCTATGTTGCTTATCTCAACATGATGCAACATCACTTCCGAGATTTGACCAATCA 1553

Db 1541 TGTGCTTAGTTTGACGCCAAACGTGGGCGACATCGTTTTCATTAATATGATATGA 1600  
Qy 1554 T 1554  
Db 1601 T 1601

RESULT 10  
US-10-424-599-126502  
; Sequence 126502, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 126502  
; LENGTH: 1872  
; TYPE: DNA  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_85238C.1  
US-10-424-599-126502

Query Match 16.0%; Score 250.8; DB 13; Length 1872;  
Beet Local Similarity 50.6%; Pred. No 1.4e-66;  
Matches 749; Conservative 0; Mismatches 707; Indels 24; Gaps 5;

Qy 90 GTTATGTGAGACCAAGAAAGATTGAAGAACTTCTTAACAAAGAAATTCAGAGTCC 149  
Db 144 GGTGTGGGTGAGCCCAAGAGATGAGAGCGCTCAAGACAGGGTATCCAAAGAA 223  
Qy 150 TCTTATATCTTCTCAATCGGAATGTTAAGAACTTTGGAATGATGTTAAAGCTTC 209  
Db 224 TTCTGACCGCTTTGATGAGACATCAGAGATGATGATGATGATGATGATGATGATGAT 283  
Qy 210 TTCTATCTATG--CCTTCTCTCAAAATTTCTCTAGAGTTCTCTTTTAAACA 266  
Db 284 ATCCAAACCTATGATCTCACTATATGATGACCCCGTGTGCTTACCTGT 343  
Qy 267 TCACTGAGAAAAATCTACGCTGCTACATTTCTGTTGGTTCGGTCCAACTTCCGGTT 326  
Db 344 TCACCACTTCCTAAATACGGAATGATGATGATGATGATGATGATGATGATGATGATGAT 403  
Qy 327 AACGTGACGATCTGATTTGATCAGAGATCTTCTAAGTCTGAGTTCTACAGAA 386  
Db 404 ATTCACTTGAATCTGCAAAATTCAGAAAGATGCTAATGATGATTTTCAAA 463  
Qy 387 GAATGAGCTACCTCTTGTGTTAAACACTTGAAGGAGATGACTTATGCTCAAG 446  
Db 464 GCCCGACATGATGACATTTCA--AGCTTCTACATCAGGGTTGCAATTAATGATGG 520  
Qy 447 TGAATAATGGGCTCATCATCAAGAAATCATTAAGCCATCTTTCATATGAGATCTTAA 506  
Db 521 GCAAGTGGCTTAAACAGAAAGATGCGAGTCCAGCATTTCAATGATGAGAAATGAA 580  
Qy 507 GTTGTGTACAGTGTGTTGAAGAGTGTGATGATGATGATGATGATGATGATGATGATGAT 566  
Db 581 GCTCTGTGATCCGATTTTGGCCAGAGTGGATGATTTGATCAGCAAAATGAGAGATT 640  
Qy 567 GTTATCAAA--AACGTGAGTGAAGTGAATGATGATGATGATGATGATGATGATGATGAT 623  
Db 641 GTTATCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 700  
Qy 624 TGAAGATTTATATGATGAACAGCTTTTGAAGTATGATGATGATGATGATGATGATGATGAT 683  
Db 701 AAGGACCTTCTGCTGTCAGGCTTTGGAAGTATGATCAAGAAAGGAAAAAATATT 760



QY 684 TCAGCTTCAAGCTCAAAAGTCTTTGTGCTGAGCTTTTCAAAAGTCTTCAATCC 743  
 DB 761 CGAATTTCAAGGAAATGATTCAGCTTCAATGACGCTTTTAAGTTGCTTCAATCC 820  
 QY 744 TGGCTATAGATTTTTCGCAAGAGGAAATTTGAAGTCTGGAAGTTAGCAAGAGAT 803  
 DB 821 AGGTACAGTTTCTGCAACGATTCACAAAGAGATGAAAGCAATTTGCAAGAAAT 880  
 QY 804 AAGGAAGTGTGTTTGAAGCTGATAGAGGGGAGAGCAAAAGCTATAGATGAGAGG 863  
 DB 881 ACAGAGATCATTTATGATATCATCAACGAAAGATTTAAAGCAATCAAGAGAGAGGCC 940  
 QY 864 GGAAGATGATGAAGAGC-----CGGCGGCGAAGATTTGTTGGATTATGAT 911  
 DB 941 TACTTACATGACTTGTAGGCACTCTTGAATCAATTTAACAAGAAATCGAAAAAG 1000  
 QY 912 TCAGGCAAGAAATGTGACGGTTGAGCACTTGTGAGAGAGTGAAGCTTTTCTTCCG 971  
 DB 1001 TAGGTGTGAGGAATGATTTAAGGAACTAGTGAAGAGTGAAGCTATTTTACTTGGC 1060  
 QY 972 CGGAAACAGACAACTTCTAATCTGCTGACGTGACGACCAATCTTGTATCCATGCAACC 1031  
 DB 1061 AGGGCAGAGAGCAAAATGCAAAATGCTGCTGACATTTGTTATTTAAGCAAGCAATCC 1120  
 QY 1032 GAGGTGCGAGCCAAAGACGTGATAGGTCTCTAGGGTCTGCGGCTCAGCTGATGCC 1091  
 DB 1121 AGATGGCAAGAAAGGCTGGGGAGAGGCTTTTCCAGTGTCCG---GAATGAAAGCC 1177  
 QY 1092 TACCAAGACCAATGCTTAAAGCTTAAAGCTTGAATGATCTTGAACGAGCTTTAAG 1151  
 DB 1178 AGATTAATGAAGATTTGCTACATAAATTTGTGATGATTTTACAGAGAGTCTAG 1237  
 QY 1152 GTTGTATCCACCAATAGTACGTAACGATTCAGCGCTTAATCGATGTAAGTGAAGAGG 1211  
 DB 1238 ATTAATCCACGAGTGTATATGTTGCTCGGTATCTTGTAAAGATCAAAAGCTTGAGA 1297  
 QY 1212 GTTCAAAATCCCATGTGCGACGAGATTTCTAATCCCAATCAGAGGCTCAATCAGCA 1271  
 DB 1298 GCTTACAAATCCCGCAGAGTGAAGTGTGTGACTGTGTCAATGCTACACAGATTA 1357  
 QY 1272 AGCCATTTGGGGTAAATGACGTGAACGAATTTCAATCCAGCTCGGTTGGCGATGAGTCC 1331  
 DB 1358 CGAGTTTGGGAGAGATGATGCGGGGAATTTCAACGAAAGATTTCTCGAAGAGATTC 1417  
 QY 1332 GCGTGTGCAAAACCCCGTTGGCTTCAATCCGTTGGCTCGGAGTTTCTCAATGAT 1391  
 DB 1418 AAAGGCAACAAAGGAGAGCTTTCCTACTGCAATTTGATGAGGCTCTCGACTCTGCAT 1477  
 QY 1392 TGGTCAAGATCTTGTACTATCTTCAAGGCAAAATGACATCGCTGTAATGATCAACGCTT 1451  
 DB 1478 AGGACAAACTTTGGCTGTGTAAGCAAAAGAGTGTGTCAATGATCTCGAACGTTT 1537  
 QY 1452 CACCTTCACTTGGCTCTTACTTATGACATGACCTACCTCTTATGTTGCTTTATCC 1511  
 DB 1538 CTCCTTCATTTTCCCTCTCTATGCTCATGCCCCCTCTTATATATCATCTTCAAGCC 1597  
 QY 1512 TCAACATGATGACCAATCACTTCCGGAATGACCAAT 1551  
 DB 1598 TGAAGGTGGGCTCATCTCTTTTACGAAATCTTAGAAT 1637

RESULT 11  
 US-09-938-842A-92  
 ; Sequence 92, Application US/0993842A  
 ; Patent No. US020160378A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Harper, Jeff  
 ; APPLICANT: Krep, Joel  
 ; APPLICANT: Wang, Xun  
 ; APPLICANT: Zhu, Tong  
 ; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING  
 ; TITLE OF INVENTION: SAME, AND METHODS OF USE

FILE REFERENCE: SRIPI300-3  
 ; CURRENT APPLICATION NUMBER: US/09/938, 842A  
 ; CURRENT FILING DATE: 2001-08-24  
 ; PRIOR APPLICATION NUMBER: US 60/227,866  
 ; PRIOR FILING DATE: 2000-08-24  
 ; PRIOR APPLICATION NUMBER: US 60/264,647  
 ; PRIOR FILING DATE: 2001-01-16  
 ; PRIOR APPLICATION NUMBER: US 60/300,111  
 ; PRIOR FILING DATE: 2001-06-22  
 ; NUMBER OF SEQ ID NOS: 5379  
 ; SEQ ID NO 92  
 ; LENGTH: 1557  
 ; TYPE: DNA  
 ; ORGANISM: Arabidopsis thaliana  
 US-09-938-842A-92

Query Match 15.4% Score 241.2; DB 9; Length 1557;  
 Best Local Similarity 49.2% Pred. No. 1.2e-63;  
 Matches 754; Conservative 0; Mismatches 768; Indels 12; Gaps 4;

QY 10 GAAAGTACAGCTGCTTCAATTCGAAAGTCTTGTCTGTCTGATCTTAAGCTTGTGA 69  
 DB 4 GAACTTAATAGCACATCAATCTTAAACATGCTTCTTACCTTGTGTTTCAAG 63  
 QY 70 ATAGGAGAGGATGCTCTGTTAAGTGAACCAAGAAAGATGAAACATTTCTCT 129  
 DB 64 ATATGAAAGCTTGTGATCTCTTGTGCTCATTTGATGCTATGATGAAAGATTCAG 123  
 QY 130 AAACAGAAATTCAGAGTCTCTTATCATTTCTTATCGGAAATGTTAAAGACTTGT 189  
 DB 124 AAACAGAAATTCAGAGCAAAAGTCTTGAACGAAACCTCAGGAGATTAAG 183  
 QY 190 GGAATGAGCTTAAAGCT---TCTTCATGCTTAAAGCTTCTCTCAATATCTCT 246  
 DB 184 AGATGAAGAAAGAGCTGACCTTGTGTCTGATCCCAATTCATGATATTTTCCCT 243  
 QY 247 AGAGTCTCTCTTTTACCATCATCGAGAAATCTACGCTGCTACATTTCTGTTGG 306  
 DB 244 CGGTGTTTCCATGATCAACATGATGCTCAATGAGATGCTCAATGCTTCTATTCGG 303  
 QY 307 TTGCTGCACTTCCGTTTAAAGCTGATCCGATCTGATTTGATCAGAGATCTTCC- 365  
 DB 304 ACTGAAACAAACCAATATATCATCTCAATATGAGCTGATGCAACAGATCTTGTG 363  
 QY 366 --TAAGTGAGTTTACAGAGAAAGATGAGCTCACTTGTGTTAAACAACTTGAAGC 423  
 DB 364 AGCAATGCTGTTTACATTTATACCGGTGAAGAGCTGAGGCTTCAATCTTCCGT 423  
 QY 424 GATGAGTATAGTCTCAAGGTAAAGGCTCATCTGAAATATATAGCCCT 483  
 DB 424 AAAGACTTCTCTTATCCAAAGTATGATGATGCTGCAACGAAATTTTAAACCT 483  
 QY 484 ACTTTCATATGAGAAATCTTAAGTGTGCTTGAACAGTGTGTTGAAGAGTGAAGT 543  
 DB 484 GCTTTCATATGATGCTGCTCAAGGCTGATGACACAGAGTGGAGATTTGACCTTACG 543  
 QY 544 ATGTTGAAGAGTGGAGAAACAGAGAAATGATGAGTGTGATCAAGATTTGAGATA 603  
 DB 544 ATGTTGAAGAGTGGAGAAACAGAGAAATGATGAGTGTGATCAAGATTTGAGATA 603  
 QY 601 TATGAGTGTTCAGATTTTGACTGAAGTGTATTTAGAAACGCTTTTGAAGTAC 660  
 DB 604 AGCAAGATTTCAATTAAGTACCGCTGATATATGAACTATCTGCTTTGAAGTAC 663  
 QY 661 TATGAGATGCTGAGCACTTTTGAAGTGTGATCAAGTCAACAAAGCTTTTGTGCTGAA 720  
 DB 664 TATGCCGAAGATGAGATTTGTGAGTCAACAAAGCTTGAAGATTTATATTAAGT 723  
 QY 721 GCTTTCAAAAGTCTCATCTGCTTATGATTTTTCGACAGAGGAAATTTGAG 780  
 DB 724 TCTCTACATGAGTTTATATCCCGAACTCAATCTTCTTACGCCCTTACCACTTAA 783  
 QY 781 TCTCGAAGTTAGCAAGAGATTAAGAGTGTGTTGAAGCTGATGA---GGGCGG 837



Mon May 24 08:10:08 2004

us-09-992-901-1.rnpb

Page 11

Db	784	CTATGGAGAGCTCCATAGAAAGTGAGAACTCCATCAAAAGAAATCAATAGATTTCAGAGCTTA	843
QY	838	AGACAAACACGCTTATAGATGAGAAAGGGAAGAAATGTAAAGGACCGGCGGAAGATTGG	897
Db	844	AAATCAAAATAGTAAGACCTATGGCTATGGGAGAGATCTTCTAGAGGATCATGTGACTGCT	903
QY	898	TTGGAGATTAAATGATTCAGGCAAAAGATGTACAGGTTCAAGACATTTGGAGAGAGTATAA	957
Db	904	GCCAAATCTAACAGATATGAGAAAGATGAAATGAAATGATGATCAATAGAGGAATCAAG	963
QY	958	AGCTTTTCTTCGCGCGGGAACAGACAACTTCTTAATCTGTGTCAGTGAAGACATCTTG	1017
Db	964	AATTTCTACTATGACAGGCAAGAACTACTTTCATTTGATTAATGACATGACATGATTA	1023
QY	1018	CTATCCATGACCCCGAGTGGCAGGCCCAAGACAGTGATGAGGTCTCAGGGTCTGCGGC	1077
Db	1024	CTAAGCTTGACACCAAGGTTGGCAAGAAACTCAGAGAAAGGTTTTTATGATATGTGAT	1083
QY	1078	TCACGTATGTCCTTACCAGAGACCAATGTCGTAAAGCTTAAACGTTGAGTATGATCTTG	1137
Db	1084	AAGATTAATAATCCAGATACAGACACTTCTTCAACTCAAAATTTGATGAACATGATGTTG	1143
QY	1138	AACAGATCTTTAAGTTGATATCCACAAATAGTAGCTACGATTGACAGCGGCTAAATCGAT	1197
Db	1144	ATGAGATGCTGTGCTGTGACGACCGGTATTAATAATTTCCGAGAGGCAACAAAGAT	1203
QY	1198	GTCGAGCTAGAGGGGTCAAAATCCATGTGGCAGGAGGTTCTTAATCCCAATCATAGGG	1257
Db	1204	ATGAAAGTAGAGACACTTGGAGATCCCAAAGGGCAACAGCATTTTCATACCTGTACTGAAG	1263
QY	1258	GTCATCATGACCAAGGCATTTGGGGTAAATGACGTGAACAAATTCACAGCTCGGTTT	1317
Db	1264	ATGACAGAGACAAAGGCCATATGGGGGGAAGAGCGCAACATTTCAATCCCTTGAGATTC	1323
QY	1318	GCGAGTAGAGTCCGCGTGTGTCGCCAAACACCCCGTTGGCTTCATACCGTTTGGCCTCGGA	1377
Db	1324	GAAGACGGCAATTTTCAAGGCCACCATTCACCAACGCTCTCCCTCCGTTCTCAATCGGA	1383
QY	1378	GTTCGTACATGATGATGTCAGATCTTGCTAATCTCAGGSCCAATATGACATCGCTGTGA	1437
Db	1384	CCAAAGCTTGATTCATGCCCCAAAACCTTGGCCATGTGCAAGCCAAAGCCGTTCTCACATAG	1443
QY	1438	ATGATCCACGCTTACACCTTTCACCTTGGCTCTTAATTATGACATGCACTTACCGTCTT	1497
Db	1444	ATCCTTCAACAGTTCCAGCTTAGGCTCTCCCCGAGATTAAGACACGCGCATGTGATCAC	1503
QY	1498	ATGTTGCTTATCTCTCAACATGTCGACCCATCA	1531
Db	1504	TTGATCTCTTCCGCAATACG3CCTTACCCTGA	1537

```

; SEQ ID NO 92
; LENGTH: 1557
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-938-842A-92

Query Match          15.4%; Score 241.2; DB 11; Length 1557;
Best Local Similarity 49.2%; Pred. No. 1.2e-63;
Matches 755; Conservative 0; Mismatches 768; Indels 12; Gaps 4

Dy      10  GAAAGTAGCAGCTGGTTCATTCCAAAGGTTCTTCTGTCTGTAAATCTTAAGTCTGTGA 69
Db      4  GAACCTATATAGACATCAATCAATCTCTTAAACATCGTTCTTCTACTCTTCTGTTTCCAA 63

Oy      70  ATATGGAAGGATATCTCTGTATATGGGAGACCAAGAAAGATTGAAGACATTTCTCT 129
Db      64  ATATGGAAGCTTTGTGATCTCTCTTTGGGTCCATTGATCTATACAAAAGATTCAAG 123

Oy      130  AAACAAAGAAATCGAGGTCCTCTTATCATTTCTTCATCGAAAATGTAAAGACTTTGT 189
Db      124  AAACAAAGAAATCTCAGACCAAGAACAAATCTTTACGGAACCTCAGCGGATTAAG 183

Oy      190  GGAATATCTTAAAGCT---TCTTCATCTCTATAGCTTTCTCTCAATATCTTCTCT 246
Db      184  AAGATTAAGAAAGAAAGCTGACCTTGTGTGTGATCCAAATTCATATGATATTTCCCT 243

Oy      247  AGAGTCTCTCTTTTATCCATCACTGAGAAAATCTACGGTGTCTACATTTCTGGTTGG 306
Db      244  CGGTGTCTTCTCAGATATCAACAAATGATGTCTCAATACGAGATACATTTCTATTCTG 303

Oy      307  TTGGTTCAACTTTCCGGTTAACGGTAGCCGATCTCTGATTTGATACAGAGATCTTCTC 365
Db      304  ACTGGAACAAACCAACTATATATCACTCAATATATAGCTACGGAACAGCTTTGTGG 363

Oy      366  --TAACTGAGGTTCTACGAGAGAGATTAAGCTCACTCTTGTGTTAAACAATTGAAGC 423
Db      364  AGCAATTCGGTTTCACTATTATACCGGTGAAGAGACTGAGGCTTCATACTTTTCGGT 423

Oy      424  GATGACATACTATGTCTCAAAGGTGAAAAATGGGCTCATCATCGAAAAATCATTTACCT 483
Db      424  AAAGGACTTTCTTTATCCAAAGGTATGATTTGATGGCCATACGAATCTTGAACCTT 483

Oy      484  ACTTTTCATATGAGAGATCTTAAATGTTGCTTTACAGTGTGTGTAAGATGTACGTAT 543
Db      484  GCTTTTCCATGATGCGCTCAAGGCCATGACACACCGAATGGAGATTCACACTTAACG 543

Oy      544  ATGATGATTAATATGTCGCATTAATTATCA--GAAAAAGGTGAATGAGTAGATGTC 600
Db      544  ATATTTGAGAGAGTGGAAGAAACAGAGAAAGAAATGATGAGTTGATCAAGATTTGATA 603

Oy      601  TATGATGCTTTCAAGTTTGTGATGAATGTTATATGTAAGACAGTTTGGAAAGTAC 660
Db      604  AGCAAGAGTTCATTAATGACCGCTGACATTAATAGCACTCTGCGTTTGGAAAGATG 663

Oy      661  TATGAAGATGTGACAGCTTTTTCGACTTCAAGCTCAACAAATGCTTTCTTGTGCTGA 720
Db      664  TATGCCAAGAGATCGAATTTGTGTAGTCACAAAACAGAGCTTGAAGATTTATTAAT 723

Oy      721  GCTTTTCAAAAAGTCTCAATCTCTGATATGATTTTTTCCGACAAAGGGAATTTGAAG 780
Db      724  TCTCTCACTAAGCTTTTCAATCCCGGAATCAATCTCTTCAAGCCTTACCAACTTAA 783

Oy      781  TCTCGAAGTTTAAACAAGAGATTAAGAAATGTTGTGAAGCTGATAGA--GCAGCGG 837
Db      784  CTATGGAGGCTTCATTAAGAAAGTGAAGAACTGATCAAAAGATCATTAATTAAGCTTA 843

Oy      838  AGACAAAGACCTATATATGAGAGAGAGAAATGTAAGAGCCGAGCGGAGAAAGATTTG 897
Db      844  AAATCAAAATGTAAAGACTTATGAGCTATGAGGCGAATCTTTTAGGGGTCAATTTGACT 903

Oy      898  TTGGGATTAATGATTCAGGCAAGATGTGACGGTTTCAGACATTTGGAGAGCTGTAAA 957
Db      904  GCGAAATCTTAAGAGATHTAGAGAAAGATGAAGATGATGATGATCATAGAGAAATGCA 963

```

QY 958 AGCTTTTCTTCCGCGGAAAAGACAACTTCTAATCTGTGACGTGACGACATCTT 1017  
Db 964 AATTTCTACTATGACAGGCGAAGAACTTCTAATCTGTGACGTGACGACATCTT 1023  
QY 1018 CTATCATGACACCGGAGTGGACGCGCAAGACGTGATAGAGTCTCTAGGCTGCGGC 1077  
Db 1024 CTAGCTTGACCAAGGTGGCAGAGAACTCAGAGAGAGGTTTATATGATGTGT 1083  
QY 1078 TCAGCTGATTCCTTACCAAGACCAATGTCTTAAAGTTAAAGTTGATGTGT 1137  
Db 1084 AAAAGTAAATCCAGATACAGACCTTCTTAACTCAATGTAAATGTGTGTG 1143  
QY 1138 AACGATCTTTAAGTTGTATACCAATAGTACATTCGACGCGCTAATCGGAT 1197  
Db 1144 ATGAGTGTCTTCTGTGTACGACCGGTGATTAATTAATCCGAGAGGCAACAGAT 1203  
QY 1198 GTGAAGCTTGAAGGATACAAATCCATGTGGACGAGCTTCTAATCCCAATCAAG 1257  
Db 1204 ATGAAGTAGACACTTGAGATCCAAAGGCAAGCAATATCTACTGTAAG 1263  
QY 1258 GTCCATCATGACCAAGCAATTTGGGTATAGAGTGAACGAATTCATCCAGCTGCTT 1317  
Db 1264 ATGACAGAGACAGGCGCATATGGGGGAGAGAGCGCGAACAATGATCTTGTAGATTC 1323  
QY 1318 GCGATGAGTCCGCGCTGTGCGAAACACCCCGTGGCTTCAATCCGTTGGCCTCGGA 1377  
Db 1324 GAAAAGGCAATTTCTCAAGCACCACTTACCCAAAGCTCTCTCCCGTTCTCAATCGGA 1383  
QY 1378 GTTCATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1437  
Db 1384 CCAAGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1443  
QY 1438 ATGATCCAGCTTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCACT 1497  
Db 1444 ATCTTCAACGTTCCAGCTTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCACT 1503  
QY 1498 ATGTTGCTTATCTTCAACATGATGATGATGATGATGATGATGATGATGATGATGAT 1531  
Db 1504 TTGATCTCTTCCGCAATAGGCTTACCGTAA 1537

RESULT 13  
US-10-425-114-15634  
/ Sequence 15634, Application US/10425114  
/ Publication No. US2004003488A1  
/ GENERAL INFORMATION:  
/ APPLICANT: Zhou, Jinsong  
/ APPLICANT: Kovalic, David K.  
/ APPLICANT: Screen, Steven E.  
/ APPLICANT: Tabaska, Jack E.  
/ APPLICANT: Cao, Yongwei  
/ TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
/ FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
/ FILE REFERENCE: 38-21(5313)B  
/ CURRENT APPLICATION NUMBER: US/10/425,114  
/ NUMBER OF SEQ ID NOS: 73128  
/ SEQ ID NO 15634  
/ LENGTH: 1574  
/ TYPE: DNA  
/ ORGANISM: Zea mays  
/ FEATURE:  
/ OTHER INFORMATION: Clone ID: LIB060-044-A8\_F11  
US-10-425-114-15634

Query Match 14.9%; Score 232.4; DB 13; Length 1574;  
Best Local Similarity 50.2%; Pred. No. 6,6e-61;  
Matches 631; Conservative 0; Mismatches 621; Indels 6; Gaps 2;  
QY 284 ACGGTGCTACATTTCTGTTGTTGGTTCGATCACTTCCGATTAACGGTAGCGATCTCT 343  
|||||

Db 114 ACGGAGGACGAGCATGCTGTTGCGGCCCAAGCCCAAGGATGACATGCTGACCGCG 173  
QY 344 ATTATGACAGAGATCTTCTTAATGCTGATGATGATGATGATGATGATGATGATGATGAT 403  
Db 174 ACGTGGCAGAGAGCTGCTGTCACAAAGTTCGCGCACTTCGAGAGGCTCAAGTCCCG 233  
QY 404 TGGTTAAACAACTTGAAGCGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 463  
Db 234 CGCTGCAAGATGCTGGGCTCCGCGCTGCGACGACGAGGCGAGAGTGGGTGAAC 293  
QY 464 ATGCAAAATCATTAATGCTTCTTCAATGATGATGATGATGATGATGATGATGATGATGAT 523  
Db 294 ACCGAGAGATCTTCAACCATGCTTCAATGATGATGATGATGATGATGATGATGATGATGAT 353  
QY 524 TGTGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 580  
Db 354 TTTCTAGCTGCTGAGAGAGTGTGTGAGAGTGTGTGAGAGTGTGTGAGAGTGTGTGAGAG 413  
QY 581 GTGAAGTGAAGT 640  
Db 414 GCTGTGAGAGT 473  
QY 641 GAAGCTTTTGAAGT 700  
Db 474 GCACTGTGTGAGAGAGT 533  
QY 701 AATGCTTCTTGT 760  
Db 534 AGCTTCTTGT 593  
QY 761 CGACAGAGAGATTTGAAGT 820  
Db 594 CTACCGGACAGACAGAGATGT 650  
QY 821 AGCTATGAG 880  
Db 651 GAGATATATGAG 710  
QY 881 CGCGCGGAGAGATTTGT 940  
Db 711 TGT 770  
QY 941 TTGTGAGAGATTTGAAGT 1000  
Db 771 TGT 830  
QY 1001 CGTGAAGACATCTTGT 1060  
Db 831 TGT 890  
QY 1061 TGT 1120  
Db 891 TGT 950  
QY 1121 CGT 1180  
Db 951 TGT 1010  
QY 1181 GACGCGTAAATCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1240  
Db 1011 GACAAAGTAAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1070  
QY 1241 TAATCCATATAGCGGTCTCATATGACCAAGCAATTTGGGTAAATGACGTGAACAT 1300  
Db 1071 AGCTCCGCTGTCTCATATGACCAAGCAATTTGGGTAAATGACGTGAACAT 1130  
QY 1301 TCAATCCAGCTCGGTTTGGAGT 1360  
Db 1131 TTAGGCTGAGAGT 1190  
QY 1361 TACGTTTGT 1420  
Db 1191 TCCGTTGT 1250

QY 1421 AATTGACATCGCTGTATGATGCAACGCTTCACTTTGACTTGGCTCTACTATACAGC 1480  
DB 1251 AGATGGCGCTCTGCTATATCTCCAGCGGTTTGAAGTTCGGGCTCGGCATCTTAGCTC 1310  
QY 1481 ATGACCTACCGTCTTATGTTGCTTATCTCTCAACATGATGACCAATACCTTCG 1538  
DB 1311 ATGCGCCGCAATACATATTAACCTTGCATCTTATGACGCTGGGCACTTAAGCTCAG 1368

## RESULT 14

US-10-424-599-74806  
; Sequence 74806, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53223) B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 74806  
; LENGTH: 4300  
; TYPE: DNA  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_38563C.1  
US-10-424-599-74806

Query Match 14.8% Score 231.4; DB 13; Length 4300;  
Best Local Similarity 51.4%; Pred. No. 2,86-60;  
Matches 678; Conservative 0; Mismatches 611; Indels 30; Gaps 5;  
QY 257 CTTTTCACATCATGAGAAAAATCTACGGTGTACATTTCTGTTGTTGGTTCGTCGA 316  
DB 1288 CTTTAAACATCATGTCATCAAAATTCGGCAAGAAATCATTTTTCGGAAGTACCA 1347  
QY 317 CTTTCGGTTAACGCTACCGATCTCTGATTGATGAGAGATCTCTTAAGTGAAGT 376  
DB 1348 AACCAAAAGTGTATCACTGACCTTGCAAAATCAAGATCTTCAACAAATCCAG 1407  
QY 377 TCTACAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 436  
DB 1408 ACTTTGAGAAACCAAGTTGAGTCTATGTC---AGTTGTGGGCTCTGGAATTGCA 1464  
QY 437 GTCTCAAGGTGAAAAATGGGCTCATCTATGAAAAATCATAGCCCTTTCATATG 496  
DB 1465 ATTGTAGAGCGAAGATGGCGTACATCGAAGATTAATACCTGCTTTCATTTAG 1524  
QY 497 AGATCTTAAGTGTGCTGATGACAGTTGTTGAAAGTGTGATGATGATGATGATGAT 556  
DB 1525 AAAAATTTGAAGTTATGTTACCAATTTCTTGAATGTTGAGAGATGATGTTCCAAAT 1584  
QY 557 GGTCCGATTAAGTTATGAGAAACGGGTGAAGTTGAGTGAATGCTATGAGTGGTTT 616  
DB 1585 GGGAAAGTTGTTGCTTCAATGCAATTCGAAATGATGATGATGATGATGATGAT 1644  
QY 617 TTTTGACTGAAGATGATTAAGTGAAGCTTTTGAAGTGAAGTGAAGTGAAGTGAAG 676  
DB 1645 ATTATACATGATATATATATCTAATACGATTTTGAAGAGATTAAGATGGAAC 1704  
QY 677 CAGTTTTCGATCTCAAGCTCAACCAATGCTTTCTTGTGCTGAAGCTTTTCAAAAGCT 736  
DB 1705 GTATATTTGACTTCTGAAGAGCAAACTGACCTTATGATGAAA---TTGAGAGAGCTT 1761  
QY 737 TCTTCCGCTATTAATTTTTCGAGCAAGAGGAAATTTGAGTCTCGAAGTTAGACA 796  
DB 1762 ACATTCAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1821

QY 797 AGGAGATTAAGAGATGCTGTTGTAAGCTGATGAGCGGCGAGACAAAACGCTATGATG 856  
DB 1822 CTGAGATTAAGAGATTAATTAAGGATGATCAACAAAGAGATGCAATGAAGGCTG 1881  
QY 857 GAGAGAGAGAGATTAAG 907  
DB 1882 GTAAAGCTTGAACAAATGATGTTGGGATGCTTTTGAATCAAAACGATGAGAAATG 1941  
QY 908 TGAATCAGCAAGAT-----GTACGAGTTTCAGACATGAGATGAGAGATGTA 955  
DB 1942 AAGATCATGAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2001  
QY 956 AAAGCTTTTCTTGGCGGAGAAACAGACATTTTATGCTGACGAGAGACATCT 1015  
DB 2002 ATGCAATCTCATAG 2061  
QY 1016 TGCTATCATGACCCGAGAGTGGCAGAGCCAAAGCAGTATGAGTCTCAAGGATCTGCG 1075  
DB 2062 TATTTGATGATATCTCTCATTTGCAAGAGAGAGAGAGAGAGAGAGATGATGAT 2121  
QY 1076 GCTCACGATGATGCTCCATCAAGAGACATGCTGTTAAGCTTAAACGTTGATGATCT 1135  
DB 2122 GTAAC---AAAGCCAGACTATATATGAGTCACTTAAATTTGACCATGATTT 2178  
QY 1136 TGAAAGATCTTAAAGTTGTATCCAAATGATGATGATGATGATGATGATGATGAT 1195  
DB 2179 TATAGAGATCTTAAAGTTGTATCCAAATGATGATGATGATGATGATGATGATGAT 2238  
QY 1196 ATGTGAAGCTAGAGAGGTATCAAAATCCATGTCAGAGAGCTTCTAATCCATATATG 1255  
DB 2239 ATGTGAAGCTAGAGAGGTATCCCTTACCTGAGAGAGAGAGAGAGAGAGAGAGAG 2298  
QY 1256 CGGTTCATCATGACCAAGCATTTGGGATGATGATGATGATGATGATGATGATGAT 1315  
DB 2299 TGATTCACAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2358  
QY 1316 TTGGGATGAGAGTGGCGGCTGTCACCAACCCCGTGGCTTCATACGCTTGGGCTCG 1375  
DB 2359 TTGCTGAAGAGATGACAAAGCAAAAGCCAAAGTTGTTTCCATTTGAGATGG 2418  
QY 1376 GAGTTGATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1435  
DB 2419 GTCTAGAGTGTGCTTGGGCAAAATTTGCTTATTAAGAGGAGATGATGATGAT 2478  
QY 1436 TAAATGATCAAGCTTCAACCTTCACTGCTCTTATTCAGATGACATGACCTACCGTCC 1495  
DB 2479 TACTCTTAAGAGCTTCTCATTTGAGCTTCTCTACATGACATGCTCCAGTCACTG 2538  
QY 1496 TTATGTTGCTTATCTCAACATGATGATGATGATGATGATGATGATGATGATGAT 1554  
DB 2539 TGCTACTCTTAATCCAAATTTTGTGCTCATATCATTTAATGATGATGATGAT 2597

## RESULT 15

US-10-425-114-24276  
; Sequence 24276, Application US/10425114  
; Publication No. US20040034888A1  
; GENERAL INFORMATION:  
; APPLICANT: Liu, Jingdong  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Screen, Steven E  
; APPLICANT: Tabaska, Jack B  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(5313) B  
; CURRENT APPLICATION NUMBER: US/10/425,114  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 73128  
; SEQ ID NO 24276  
; LENGTH: 1808  
; TYPE: DNA

Mon May 24 08:10:08 2004

us-09-992-901-1.rnpb

Page 14

ORGANISM: Zea mays  
FEATURE:  
OTHER INFORMATION: Clone ID: LIB3608-017-B5\_FLI  
US-10-425-114-24276

Query Match 14.7% Score 230.2; DB 13; Length 1808;  
Best Local Similarity 50.2%; Pred. No. 3.6e-60;  
Matches 641; Conservative 0; Mismatches 613; Indels 23; Gaps 2;

QY 284 ACGGCTCACTTTGTTGGTTCGCACTTTCCGTTACGAGTGAAGTCACTT 343  
DB 343 ACGGCAAGACGTCTTCACTGGTTCGCTCCGCTCCCAAGGAGTGCCTGACCCG 402  
QY 344 ATTGATCAAGAGATCTTCTTAAGTGAAGTTTACGAGAAGTGAAGTCACTT 403  
DB 403 ACTGAGGAGGAGCTGATGCTCAAGATTCGCGCACTTGCAGAGGCCCAAGTCCCG 462  
QY 404 TGGTTAAACACTTGAAGCGATGACCTTATGCTCAAGAGTGAAGTGGCTCATC 463  
DB 463 CGCTCAACCAAGCTGTTCTCCAGCGTGTGCCCAACCAAGAGGCGAAGTGGTCAAGC 522  
QY 464 ATGGAATAATCATTAACCTTACTTTTCAATGAGAATCTTAAGTGTGACCAATG 523  
DB 523 ACAGAGAGATCCCTCAACCTTCCTTCACCTGAGAGCTCAAGCTCATGCTGCGGCT 582  
QY 524 TGTGAAGAGTGTGACTGATATGTTGATTAATGTCGCAATTAAGTTATCAGAAA 583  
DB 583 TCTCTGCTGCTGTGAAGAGCTGTGAGCGGATGGCGGAGTGGCTTGGCCCGACGCT 642  
QY 584 AAGTTAGTAGTAGTCTATGAGTGTTCAGATTTTGAATGAATGTTATGATGAA 643  
DB 643 CGTGGAGCTGAGCGTGAACCCGAGCTTCAACCTCCAGCGATGTCATCTTCGCA 702  
QY 644 CAGCTTTGGAAGTGAATGAAGATGTCGAGCAATTTTCACTCAAGCTCAACAA 703  
DB 703 CCGGCTTCGAGCAAGTACTACTGAGAGAGAAAGATTTTCCAGCTCCAGGCGAGCA 762  
QY 704 TGCTTTTGTGCTGAAGCTTTTCAAAAAGCTTCACTTCTGCTATGATTTTCCGA 763  
DB 763 CTGAGCGCTCATGCTCATCATGACAAATTCGCCCTTCGAGTACATGCTTGCTTA 822  
QY 764 CAAGAGGGAATTTGAAGTCTCGAAGTTAGACAAGGATTAAGAAAGTGTGTAAGC 823  
DB 823 CGAATAACCAACCAAGATGCTGCAATTAACCAAGCAATGCACTGATCTTCGGGGCC 882  
QY 824 TGATGAGCGCGGAGACAAACGCTATGATGAGAGAGGAGAA----- 870  
DB 883 TGATCGGTAAAGATGCAAGCTATGAAACAGGTGAGCGCAAAAGATGACCTGTGG 942  
QY 871 -----TGTAAGGAGCGGCGGAGAGATTTGTGGATTAATGATTCAGGCAAG 921  
DB 943 GCTTACTGCTGAGTCAACGCGAGAGACAGGCAATCAGAGCGGCCAGCCGCGCAAG 1002  
QY 922 AATGTGACGTTTCAGACATGTGAGAGAGTGAAGCTTTTCTTCCCGGAAACAG 981  
DB 1003 GAT-TGAGATGAGAAAGTCAATGAGAGTGCAGAGCTGTTTACTTCCGCGAAATGAG 1061  
QY 982 ACAACTTCTAATCTGCTGAGCGTGAACACATCTTGCTATCCATGACCCGAGTGGCAG 1041  
DB 1062 ACGAGCTGCTGCTGCTCAGTGCATGCTCTGCTAAGCATGCAACCGAGATGCGAG 1121  
QY 1042 GCCAAGACGCTGATGAGGCTCTCAGGCTCTCGGCTCAGTATGCTTACCAAGAGAC 1101  
DB 1122 GACCGTGAAGAGAGAGTCTTCGACTGTTGGGAAAGAAACCAACCGGAGATGAGAC 1181  
QY 1102 CATGCTTAAAGCTTAAACGTTGATGATGATCTTGAACGAGCTTTAAGGTTGATCA 1161  
DB 1182 GAGCTGAGTCCGCTCAAAACAGTACCATGATCTTACAGAGTCTCCGCGTGTACCG 1241  
QY 1162 CCAATAGTAGCTAGATGACGCGCTAAATCGATGTGAAGCTAGAGAGGTACAAATC 1221  
DB 1242 CCGGCAATCGCTTCAGAGAGAAACGATCAAGAGATGTTGTTGAGAGCTCACTGAC 1301

QY 1222 CCATGTGACAGGAGCTTCTATCCATCAATGAGCGGTCCATGATGACCAAGCCATTGG 1281  
DB 1302 CTTGCGGCGGAGACCTGTGAGAGCTGCCGTGTGTTATCCACACGACCCCGACATCTGG 1361  
QY 1282 GGTATGACGTGAACGAATTTCAATCCAGCTCGTTTTCGAGTGAAGTGGCGGCTGCGC 1341  
DB 1362 GGAACGATGCGCACAGAGTTGAGCGGAGAGTTCGCGAGGGGAGTGGCAGAGCTCC 1421  
QY 1342 AAACACCCGTTGCTTCAATCCGTTTGGCTTCGAGTTGTACATGATGTCAGAT 1401  
DB 1422 AAGATAGAGTAAAGGTTCTTCCCTTGGGCTGGGAGCGGAGATCTGATCGGCGAAG 1481  
QY 1402 CTGCTATATCTTACGCGCAATTTGACACTGCTATATGATTCACAGCTTCACTTCA 1461  
DB 1482 TTGCGCTTGTGAGGCGCAAGATGCGCTGAGCATGATCTTCAAGCGCTTCCAGTTGAG 1541  
QY 1462 TTGCTCTATCTTATCAGATGACACTACGCTCTTATGTTGCTTTATCTTCAACATGT 1521  
DB 1542 CTGCGCGGAGCTATACCATGTTGCGGCGAGTTATTAATGCTGCGCCGATGATGT 1601  
QY 1522 GCACCAATCACCTTCG 1538  
DB 1602 GCCCAGATCAAGCTCAG 1618

Search completed: May 22, 2004, 22:05:02  
Job time : 743 sec

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OW protein - protein search, using sw model

Run on: May 19, 2004, 10:17:15 ; Search time 59 Seconds  
(without alignments)  
2490.249 Million cell updates/sec

Title: US-09-992-901-2

Perfect score: 2730  
Sequence: 1 MEBSSSWFIPKVLVLSVL.....LYPQHGAETFRRLTNHED 520

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :  
1: Geneseq\_29Jan04:\*  
2: Geneseq\_1980s:\*  
3: Geneseq\_1990s:\*  
4: Geneseq\_2000s:\*  
5: Geneseq\_2002s:\*  
6: Geneseq\_2003as:\*  
7: Geneseq\_2003bs:\*  
8: Geneseq\_2004s:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2730	100.0	520	3 AAB23917	Aab23917 Arabidops
2	1163.5	42.6	527	5 ABB08079	Abb08079 Maize cyt
3	1098	40.2	512	5 AAU97096	Aau97096 Abscisic
4	1089.5	39.9	512	5 AAU97097	Aau97097 Abscisic
5	943	34.5	518	3 AAG31392	Aag31392 Arabidops
6	921.5	33.8	454	6 ADA48312	Ada48312 Rice prot
7	919.5	33.7	458	6 ADA48676	Ada48676 Rice prot
8	917	33.6	484	3 AAG31393	Aag31393 Arabidops
9	864.5	31.7	456	3 AAG31394	Aag31394 Arabidops
10	543.5	19.9	505	5 ABU05536	Abu05536 Breast ca
11	543.5	19.9	505	5 AAE14447	Aae14447 Human dru
12	540.5	19.8	505	5 AAU83606	Aau83606 Human PRO
13	540.5	19.8	505	5 ABU10221	Abu10221 Human cyt
14	540.5	19.8	505	6 ABU80753	Abu80753 Human PRO
15	540.5	19.8	505	6 ABO33719	AbO33719 Novel hum
16	540.5	19.8	505	6 ABU82062	Abu82062 Novel hum
17	540.5	19.8	505	6 ABU72242	Abj72242 Human PRO
18	540.5	19.8	505	6 ABU72370	Abj72370 Human PRO
19	540.5	19.8	505	6 ABO34265	AbO34265 Human sec
20	540.5	19.8	505	6 ABJ72072	Abj72072 Human mem
21	540.5	19.8	505	7 ADB83520	Adb83520 Novel hum
22	540.5	19.8	505	7 ADB80626	Adb80626 Novel hum
23	540.5	19.8	505	7 ABB73167	Abb73167 Novel hum
24	540.5	19.8	505	7 ABB78249	Abb78249 Novel hum
25	540.5	19.8	505	7 ADB84897	Adb84897 Human PRO

26	540.5	19.8	505	7 ADB78003	Adb78003 Novel hum
27	540.5	19.8	505	7 ADB87069	Adb87069 Human PRO
28	540.5	19.8	505	7 ADB84651	Adb84651 Human PRO
29	540.5	19.8	505	7 ADB83766	Adb83766 Novel hum
30	540.5	19.8	505	7 ADB872921	Adb872921 Novel hum
31	540.5	19.8	505	7 ADC36759	Adc36759 Human PRO
32	540.5	19.8	505	7 ADC21749	Adc21749 Human PRO
33	540.5	19.8	505	7 AA839879	Aae39879 Human cyt
34	540.5	19.8	505	7 ADC49780	Adc49780 Novel hum
35	540.5	19.8	505	7 ADC48979	Adc48979 Novel hum
36	540.5	19.8	505	7 ADC49496	Adc49496 Novel hum
37	540.5	19.8	505	7 ADC47957	Adc47957 Novel hum
38	540.5	19.8	505	7 ADC47102	Adc47102 Novel hum
39	540.5	19.8	505	7 ADC77977	Adc77977 Novel hum
40	540.5	19.8	505	7 ADD06212	Add06212 Novel hum
41	540.5	19.8	505	7 ADC77731	Adc77731 Novel hum
42	540.5	19.8	505	7 ADD50694	Add50694 Novel hum
43	540.5	19.8	505	7 ADD50940	Add50940 Novel hum
44	540.5	19.8	505	7 ADD50421	Add50421 Human PRO
45	540.5	19.8	505	7 ADD50175	Add50175 Human PRO

ALIGNMENTS

RESULT 1  
AAB23917  
ID AAB23917 standard; protein; 520 AA.  
XX  
AC AAB23917;  
XX  
DT 18-JAN-2001 (first entry)  
XX  
DE Arabidopsis thaliana bas1 protein sequence SEQ ID NO:2.  
XX  
DE Arabidopsis thaliana; bas1; promoter; cytochrome P450; CYP72B1; plant;  
XX  
KW brassinosteroid signalling; brassinosteroid synthesis; brassinolide.  
XX  
OS Arabidopsis thaliana.  
XX  
PN WQ200053302-A2.  
XX  
PD 21-SEP-2000.  
XX  
PF 16-MAR-2000; 2000WC-US006915.  
XX  
PR 16-MAR-1999; 99US-0124570P.  
XX  
PR 14-DEC-1999; 99US-0170931P.  
XX  
PR 20-DEC-1999; 99US-0172832P.  
XX  
RA (SALK ) SALK INST BIOLOGICAL STUDIES.  
XX  
PI Neff MW, Chory J;  
XX  
DR WPI; 2000-638195/61.  
XX  
DR N-PSDB; AAA99072.  
XX  
PT Transgenic plants having modulated brassinolide synthesis resulting in  
XX  
PT insect resistance, dwarfism and darker-green foliage compared with wild-  
XX  
PT type plants, have nucleic acid encoding BAS1 polypeptide in its genome.  
XX  
PS Disclosure; Fig 1B; 104pp; English.  
XX  
CC The present invention describes a genetically modified plant (I)  
CC comprising at least one exogenous nucleic acid sequence encoding a BAS1  
CC polypeptide, homologue or functional fragment, in its genome or at least  
CC one regulatory sequence that modulated expression of endogenous bas1 gene,  
CC homologue or functional fragment, and which is characterised as having  
CC modulated brassinolide activity or synthesis. The bas1 gene encodes a  
CC cytochrome P450 (CYP72B1), which has a role in brassinosteroid signalling  
CC or synthesis. Overexpression of the bas1 gene in plants causes a dark  
CC green, dwarf phenotype which mimics plants that have low levels of the  
CC plant hormone, brassinolide. Overexpression of the bas1 gene also

Page 2

XX  
PA (KENT ) UNIV KENTUCKY RES FOUND.  
XX  
XX  
PI Barrett M;  
XX  
WPI: 2002-470227/50.  
RR

DR N-PSDB; ABL60758.

PT Nucleic acid molecules encoding a Zea mays cytochrome P450 CYP72A1, useful for the production of plants resistant to heterocyclic,

claim 13. Fig 1: 21nm; English.  
 PT sulfonylurea and substituted urea herbicides and organophosphate  
 insecticides e.g. Classic and Pursuit.  
 PT  
 XX

XX The invention relates to isolated nucleic acid molecules encoding Zea  
CC cytochrome P450 designated CYP72A1 and CYP92A1. The  
CC mays (maize) polynucleotides can be expressed by standard recombinant methodology. The

CC nucleic acids may be used in genetic engineering protocols to transduce  
CC plants and other eukaryotes e.g. yeast, maize (especially), soybean,  
CC beet, tobacco, wheat, barley, poppy, rape, sunflower, alfalfa, sorghum,  
CC carrot, tomato, lettuce, chive, pepper, melon

CC and cabbage. They may be used in this way to confer protection against  
CC heterocyclic herbicides, sulfonylurea herbicides, substituted urea  
CC herbicides and/or organophosphate insecticides. The present sequence

XX Sequence 527 AA;  
SQ

Best Local Similarity 42.5%; Pred. No. 1.le-104, 5  
Matches 224; Conservative 103; Mismatches 181; Indels 19; Gaps 5

```

      8 MUREASPEMALAGA-VASVSYSLMVLVAMTLEWAMWTPWRILDRALRAQGLNGRTRYRLEFTGDLR 66

```

67 ETARVNRERARKKCLPLGCHDITPRVQPMHSHSTIEYKUSFTWIGOPTPRVXIIPPELVKE 126

127 VLSNKGFGHGKPRSSRIGRLIANGLVNHGDEKMAKRRILNPAFHEKIKGMPEVSTCC 186

```

187 IEMTRMDSMSSGSSSEIDWPEFQNLGDIVISRTAFSGSYQZERRIFELQGLAERLI 246

```

[illegible]

Db 301 LGLLLESNTROSNGNASTGLTTEDVIEECKLFYFAGMETISVLLTWTLLVLSMHPWQQR 360

Db 361 AREEVLSHFG-RTPPDYDSLSRLKIVTMILHEVLRLYPEALFLTRRTYKNEHLGGIKYPA 419

Db 420 GVETLLPVFIHHDPDIWCKDASEFNPERFANGISATRHQAFFFGGGERICIGQSA 479

480 LLAAVTTCTTQQFSEFLSPSYTHAPYVITTHPQGAQIRLKKS 526

RESULT 3
AAU97096
ID   AAU97096 standard; protein; 512 AA.

Page 2

Page 21,

Page 2

Page 2

```

Page 2
1,
Zealand. The
nsform
an,
rgan,
r, melon
alnst
nce
ape 5
GNVX 60
|:::
GDDR 66

```

Page 2

1,  
Zea  
The  
nsform  
an,  
rghum,  
r, melon  
ainst  
ance

aps 5  
NNVK 60  
|:::  
GGDRR 66  
lure 119  
|:::  
TIVE 126

```

1,
Zea
    . The
    meform
    at,
    ighum,
    r, melon
    alnst
    a
    ance
ape 5
GNVK 60
|:::
GGDJR 66
|:::
LIRE 119
|:::
LVIRE 126
LRSV 179
STOC 186

```

```

1,
Zea
  . The
  msform
  an.
  rghum,
  r, melon
  alnst
  a
  nce
ape 5
GNVK 60
msform
GGDIR 66
LIRE 119
LIVE 126
LRSV 179
STCC 186
LCA 239
ERI 246

```

```

1,
Zea
    . The
    nsform
    an,
    rghum,
    r, melon
    alist
    a
    nce
    eaps 5
    GNYK 60
    ::::
    GGDUR 66
    LLIRE 119
    LLVRE 126
    LKSV 179
    STCC 186
    LICA 239
    ERLI 246
    AKRL 299
    KDDL 300

```

1	Zeal	5
2	Zeal	5
3	Zeal	5
4	Zeal	5
5	Zeal	5
6	Zeal	5
7	Zeal	5
8	Zeal	5
9	Zeal	5
10	Zeal	5
11	Zeal	5
12	Zeal	5
13	Zeal	5
14	Zeal	5
15	Zeal	5
16	Zeal	5
17	Zeal	5
18	Zeal	5
19	Zeal	5
20	Zeal	5
21	Zeal	5
22	Zeal	5
23	Zeal	5
24	Zeal	5
25	Zeal	5
26	Zeal	5
27	Zeal	5
28	Zeal	5
29	Zeal	5
30	Zeal	5
31	Zeal	5
32	Zeal	5
33	Zeal	5
34	Zeal	5
35	Zeal	5
36	Zeal	5
37	Zeal	5
38	Zeal	5
39	Zeal	5
40	Zeal	5
41	Zeal	5
42	Zeal	5
43	Zeal	5
44	Zeal	5
45	Zeal	5
46	Zeal	5
47	Zeal	5
48	Zeal	5
49	Zeal	5
50	Zeal	5
51	Zeal	5
52	Zeal	5
53	Zeal	5
54	Zeal	5
55	Zeal	5
56	Zeal	5
57	Zeal	5
58	Zeal	5
59	Zeal	5
60	Zeal	5
61	Zeal	5
62	Zeal	5
63	Zeal	5
64	Zeal	5
65	Zeal	5
66	Zeal	5
67	Zeal	5
68	Zeal	5
69	Zeal	5
70	Zeal	5
71	Zeal	5
72	Zeal	5
73	Zeal	5
74	Zeal	5
75	Zeal	5
76	Zeal	5
77	Zeal	5
78	Zeal	5
79	Zeal	5
80	Zeal	5
81	Zeal	5
82	Zeal	5
83	Zeal	5
84	Zeal	5
85	Zeal	5
86	Zeal	5
87	Zeal	5
88	Zeal	5
89	Zeal	5
90	Zeal	5
91	Zeal	5
92	Zeal	5
93	Zeal	5
94	Zeal	5
95	Zeal	5
96	Zeal	5
97	Zeal	5
98	Zeal	5
99	Zeal	5
100	Zeal	5

[illegible]

Word	Count
Zeal	5
transform	60
transform	66
transform	119
transform	126
transform	179
transform	186
transform	239
transform	246
transform	299
transform	300
transform	349
transform	360
transform	409
transform	419
transform	469
transform	479

	Zea	
	. The	
	nsform	
	an,	
	rghum,	
	i melon	
	ainst	
	a	
	nce	
	aps	5
	GNVX	60
	:::	
	GDLR	66
	LIRE	119
	:::	
	LVRK	126
	LKSV	179
	SSTC	186
	LLOA	239
	ERDI	246
	LAKD	299
	KKDL	300
	WQAK	349
	MOBR	360
	KIPC	409
	KYPA	419
	ONTA	469
	QSFA	479

	Zea	
	. The	
	perform	
	an,	
	rchum,	
	r, melon	
	ainst	
	a	
	nance	
	aps	5
	GNVK	60
	:::	
	GJDR	66
	LINE	119
	:::	
	LIVE	126
	LSV	179
	STCC	186
	LLCA	239
	ERI	246
	AKDL	299
	:::	
	KDDL	300
	WOKR	349
	:	
	WER	360
	KIPC	409
	KYPA	419
	ONLA	469
	:::	
	SQSPA	479

Page 2	1.	Zea	The transform an, rchum, r melon ainst a nce	aps 5	GNVK 60 ::: GDDR 66	LINE 119 ::: LVEE 126	LKSV 179	STOC 186	LLCA 239	ERLI 246	SXDL 299 ::: KKDL 300	MMOK 349	MOER 360	KKPC 409	KKYP 419	ONLA 469	OSFA 479
--------	----	-----	---	-------	---------------------------	-----------------------------	----------	----------	----------	----------	-----------------------------	----------	----------	----------	----------	----------	----------

XX AC AAU97096;  
 XX DT 24-SEP-2002 (first entry)  
 XX DE Abscisic acid (ABA) hydroxylase CYP72A7.  
 XX KM ABA hydroxylase; abscisic acid hydroxylase; transgenic; plant; CYP72A7;  
 XX KM enzyme.  
 XX OS Arabidopsis thaliana.  
 XX PN WO200246377-A2.  
 XX PD 13-JUN-2002.  
 XX PF 06-DEC-2001; 2001WO-CA001756.  
 XX PR 07-DEC-2000; 2000US-0251518P.  
 XX PA (CANA) NAT RES COUNCIL CANADA.  
 XX PI Krochko JE, Cutler AJ, Abrams SR;  
 XX PI MPI: 2002-519663/55.  
 XX DR N-PSDB; ABR50985.  
 XX PT New isolated and purified DNA that encodes protein having abscisic acid  
 PT (ABA) hydroxylase activity, useful for altering catabolism of abscisic  
 PT acid in plants.  
 XX BS Disclosure; Page 106-108; 117pp; English.  
 CC The invention relates to an isolated and purified DNA (I) that encodes a  
 CC protein having abscisic acid (ABA) hydroxylase activity. (I) is useful  
 CC for producing a transgenic plant which involves introducing (I) into a  
 CC genome of the plant or its part, and carrying out plant growth and  
 CC development. (I) is useful for modifying catabolism of ABA or ABA  
 CC analogues in a plant. ABA hydroxylase clone pBR10-30-3 cDNA sequence is  
 CC useful for identification of related sequences from other plant species.  
 CC (I) is also useful for altering ABA and ABA analog catabolism in plants.  
 CC The present sequence represents abscisic acid (ABA) hydroxylase CYP72A7  
 CC  
 XX SQ Sequence 512 AA;  
 Query Match 40.2%; Score 1098; DB 5; Length 512;  
 Best Local Similarity 40.2%; Pred. No. 2,8e-98;  
 Matches 208; Conservative 117; Mismatches 166; Indels 26; Gaps 7;  
 QY 17 SVILSVIVKMSLW-WR-----PRKIEHRSKQIGRPYHFGTGNVKELVOML 67  
 DB 4 SYVAALPVLVAVVAVLWTRIVKMWIKPMLESSIKRQGLGTGYTPPLVGGIKENVMM 63  
 QY 68 KASSHPMPSSHILPRVLSFYHMKIKYGATFLVFGPTFRLTVADPDLLEIRSKSEFY 127  
 DB 64 EARSKEINVTDDITPRLPLALKMLNSGKTFITIGPLPTTVINPQIKSEVKNKXNDP 123  
 QY 128 EKNEAHPVLVKGEGDLISLKGKNAHHRKISPTFHEHNTKLVPVVLKSVTDMVDKMS 187  
 DB 124 EKASTFPLRLLAG-GLASFKDKNASHRRINPAFHEKIKMMPAFYHCSSHVCOM- 181  
 QY 188 DKLSNGE--VEVDVVEWFOILTEDYISTAFGSSYEDGRAVFLQAOQMLLCEAFOKV 245  
 DB 182 EKFLDKESPLEVDVWPMVMTADVISHTAFGSSYKGGQRIPOGLSLAELIIOAFKKS 241  
 QY 246 FIVGYRFPPTGRLKSRKLDKEIRKSLKLIERBRONAIDEGECKEPAKDLIGMTI- 304  
 DB 242 YIGSRFPYTKSRBRKAIDREVDVLRGIVSKRE-----KARAGSPANDDLLGILLE 295  
 QY 305 -----QANKVTYQDIVEECKSFFFGKQTTSNLTWTITLLSMPEQAKARDEVLAVC 358  
 DB 296 SNSESGQNGMSYEDVWKECKLFYFAGQETTSVLLVMTVLLSHHODQAPARREVMQVL 355

QY 359 GSRDVPFKDHVVKLTLSMLINESLTYPEPIVATIRRAKSDYKLGKYPGCTELLPII 418  
 DB 356 GENKKPDMESLNLKMTMTMFNEVLLRPPVAGLKEVNNKEMKGLTLPAGIQYLPFI 415  
 QY 419 AVHDDQAIWGNVDVNEFNPAFADGVPRAAKHVPVGFIPBGLGYRTICGQYLLIOLKTLTA 478  
 DB 416 LVGRDTELMGDDADAFKPERFRDLSKATQNVSPFPFGWGRICIGQNFAMLEAKMAYA 475  
 QY 479 VNIORFTEHLPATYOHAPTVLMLLTPDHGAPITFRRL 515  
 DB 476 LILQKSFELSPSTVTHAPQTVMTTRPQFQAHLLIHL 512  
 RESULT 4  
 ID AAU97097 standard; protein; 512 AA.  
 XX AC AAU97097;  
 XX DT 24-SEP-2002 (first entry)  
 XX DE Abscisic acid (ABA) hydroxylase CYP72A14.  
 XX KM ABA hydroxylase; abscisic acid hydroxylase; transgenic; plant; CYP72A14;  
 XX KM enzyme.  
 XX OS Arabidopsis thaliana.  
 XX PN WO200246377-A2.  
 XX PD 13-JUN-2002.  
 XX PF 06-DEC-2001; 2001WO-CA001756.  
 XX PR 07-DEC-2000; 2000US-0251518P.  
 XX PA (CANA) NAT RES COUNCIL CANADA.  
 XX PI Krochko JE, Cutler AJ, Abrams SR;  
 XX PI MPI: 2002-519663/55.  
 XX DR N-PSDB; ABR50986.  
 XX PT New isolated and purified DNA that encodes protein having abscisic acid  
 PT (ABA) hydroxylase activity, useful for altering catabolism of abscisic  
 PT acid in plants.  
 XX BS Disclosure; Page 108-111; 117pp; English.  
 CC The invention relates to an isolated and purified DNA (I) that encodes a  
 CC protein having abscisic acid (ABA) hydroxylase activity. (I) is useful  
 CC for producing a transgenic plant which involves introducing (I) into a  
 CC genome of the plant or its part, and carrying out plant growth and  
 CC development. (I) is useful for modifying catabolism of ABA or ABA  
 CC analogues in a plant. ABA hydroxylase clone pBR10-30-3 cDNA sequence is  
 CC useful for identification of related sequences from other plant species.  
 CC (I) is also useful for altering ABA and ABA analog catabolism in plants.  
 CC The present sequence represents abscisic acid (ABA) hydroxylase CYP72A14  
 CC  
 XX SQ Sequence 512 AA;  
 Query Match 39.9%; Score 1089.5; DB 5; Length 512;  
 Best Local Similarity 40.7%; Pred. No. 1.9e-97;  
 Matches 211; Conservative 114; Mismatches 171; Indels 23; Gaps 6;  
 QY 12 KYVLVSVLSVIVY-----KQSLWMPRKIEHRSKQIGRPYHFGTGNVKELVG 64  
 DB 2 EISVSSVYFSLAVVAVVSMWMTLKWVFTPKLERSLRQGLSTSTPILGPKKMS 61  
 QY 65 MMKASHPMPSSHILPRVLSFYHMKIKYGATFLVFGPTFRLTVADPDLLEIRSKS 124  
 DB 62 MFLAITSKRIRKTDITFRVWHP-QLMKTGRITLWFGIPITTIMDPQIKSEVFNKV 121

QY 125 EPEYKNAHPIYKQJEGDGLSLKGEKNAHHRKIIISPTHEMNILKLVVYVLSKVTMDV 184  
DB 122 YDFQKHTFFLSKII-GTGLVSYDDKMAQHRRILINPAFHEKTKMWHVPHSCSELVG 180  
QY 185 KMSDKLSENG-EVEVDVVEWFOILTEDVYSRTAFGSSYEDGRAVFRLOAQOMLLCAEAFQ 243  
DB 181 EMDKLVSDKSSGEVDVWFGILSMTADVISRTAFGSSYREGHRIFELOAEIQAQVMQAFQ 240  
QY 244 KVFIPGIRPPTPRGNLKSRLKXETRKSLKILERRRONALDGESEKKEPAAKDLIGLM 303  
DB 241 KFFIPGYIYLPFKGNRMKTAREIODILRGINKRR-ARSEGEAPSSDLIGIL 294  
QY 304 IQA-----KNTVODIVEECKSPFFACKQTTSMILMTWITILSMHEWQAKADEVLR 356  
DB 295 LBSNLQDTGNGNSTEDWMECKLPYLAGQETTSVLVWTVLISQHDWQARAREVYKQ 354  
QY 357 VCGSRDVPYTKDHYVVKITLSMTLINESLRILYPIVATIRASDVVLGGYKIPCTELLIP 416  
DB 355 VFQDKQ-PDTGEGNQLKWTMTLYEVLRLYPEVQVQLTRAIHEMKLGDITLPGVQISLP 413  
QY 417 IIVVHDDQIMGNDVNEFPAPFADGVPRAGHPVGFIPFGIGVTCIGOMLAILQAKLT 476  
DB 414 VLVVHRDTLWNGDAGEFKPERFKDJSKATNNOVSFPFPAMGRICIGQNFITLAKMA 473  
QY 477 LAVMIQRTFHLAPYQAHAPVYMLLYPQHGAPITFRRL 515  
DB 474 MSLLQRFSELPSSYVHAPYITITLYPQGAHMLMLHL 512

RESULT 5  
AAG31392  
ID AAG31392 standard; protein; 518 AA.  
XX  
AC AAG31392;  
XX  
DT 17-OCT-2000 (first entry)  
XX  
DE Arabidopsis thaliana protein fragment SEQ ID NO: 37691.  
XX  
KM Protein identification; signal transduction pathway; metabolic pathway;  
KM hybridisation assay; genetic mapping; gene expression control; promoter;  
KM termination sequence.  
XX  
OS Arabidopsis thaliana.  
XX  
PN EP1033405-A2.  
XX  
PD 06-SEP-2000.  
XX  
PF 25-FEB-2000; 2000EP-00301439.  
XX  
XX 25-FEB-1999; 99US-0121825P.  
PR 05-MAR-1999; 99US-0123180P.  
PR 09-MAR-1999; 99US-0123548P.  
PR 23-MAR-1999; 99US-0125788P.  
PR 25-MAR-1999; 99US-0126264P.  
PR 29-MAR-1999; 99US-0126785P.  
PR 01-APR-1999; 99US-0127442P.  
PR 06-APR-1999; 99US-0128234P.  
PR 08-APR-1999; 99US-0128714P.  
PR 16-APR-1999; 99US-0129845P.  
PR 19-APR-1999; 99US-0130077P.  
PR 21-APR-1999; 99US-0130449P.  
PR 23-APR-1999; 99US-0130510P.  
PR 23-APR-1999; 99US-0130891P.  
PR 28-APR-1999; 99US-0131449P.  
PR 30-APR-1999; 99US-0132048P.  
PR 04-MAY-1999; 99US-0132407P.  
PR 05-MAY-1999; 99US-0132484P.  
PR 06-MAY-1999; 99US-0132486P.  
PR 07-MAY-1999; 99US-0132487P.  
PR 07-MAY-1999; 99US-0132863P.

PR 11-MAY-1999; 99US-0134256P.  
PR 14-MAY-1999; 99US-0134218P.  
PR 14-MAY-1999; 99US-0134219P.  
PR 14-MAY-1999; 99US-0134221P.  
PR 14-MAY-1999; 99US-0134370P.  
PR 18-MAY-1999; 99US-0134768P.  
PR 19-MAY-1999; 99US-0134941P.  
PR 20-MAY-1999; 99US-0135124P.  
PR 21-MAY-1999; 99US-0135353P.  
PR 24-MAY-1999; 99US-0135629P.  
PR 25-MAY-1999; 99US-0136021P.  
PR 27-MAY-1999; 99US-0136329P.  
PR 28-MAY-1999; 99US-0136782P.  
PR 01-JUN-1999; 99US-0137222P.  
PR 03-JUN-1999; 99US-0137528P.  
PR 04-JUN-1999; 99US-0137502P.  
PR 07-JUN-1999; 99US-0137724P.  
PR 08-JUN-1999; 99US-0138094P.  
PR 10-JUN-1999; 99US-0138540P.  
PR 10-JUN-1999; 99US-0138847P.  
PR 14-JUN-1999; 99US-0139119P.  
PR 16-JUN-1999; 99US-0139452P.  
PR 16-JUN-1999; 99US-0139453P.  
PR 17-JUN-1999; 99US-0139492P.  
PR 18-JUN-1999; 99US-0139454P.  
PR 18-JUN-1999; 99US-0139455P.  
PR 18-JUN-1999; 99US-0139456P.  
PR 18-JUN-1999; 99US-0139457P.  
PR 18-JUN-1999; 99US-0139458P.  
PR 18-JUN-1999; 99US-0139459P.  
PR 18-JUN-1999; 99US-0139460P.  
PR 18-JUN-1999; 99US-0139461P.  
PR 18-JUN-1999; 99US-0139462P.  
PR 18-JUN-1999; 99US-0139463P.  
PR 18-JUN-1999; 99US-0139750P.  
PR 18-JUN-1999; 99US-0139817P.  
PR 21-JUN-1999; 99US-0139817P.  
PR 22-JUN-1999; 99US-0139899P.  
PR 23-JUN-1999; 99US-0140353P.  
PR 23-JUN-1999; 99US-0140354P.  
PR 23-JUN-1999; 99US-0140695P.  
PR 28-JUN-1999; 99US-0140823P.  
PR 28-JUN-1999; 99US-0140991P.  
PR 30-JUN-1999; 99US-0141287P.  
PR 01-JUL-1999; 99US-0141842P.  
PR 02-JUL-1999; 99US-0142154P.  
PR 06-JUL-1999; 99US-0142055P.  
PR 08-JUL-1999; 99US-0142390P.  
PR 09-JUL-1999; 99US-0142803P.  
PR 09-JUL-1999; 99US-0142920P.  
PR 12-JUL-1999; 99US-0142977P.  
PR 13-JUL-1999; 99US-0143542P.  
PR 14-JUL-1999; 99US-0143624P.  
PR 15-JUL-1999; 99US-0144005P.  
PR 16-JUL-1999; 99US-0144085P.  
PR 16-JUL-1999; 99US-0144086P.  
PR 16-JUL-1999; 99US-0144325P.  
PR 19-JUL-1999; 99US-0144331P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 19-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144352P.  
PR 20-JUL-1999; 99US-0144632P.  
PR 21-JUL-1999; 99US-0144884P.  
PR 21-JUL-1999; 99US-0144814P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 22-JUL-1999; 99US-0145085P.  
PR 22-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145089P.  
PR 22-JUL-1999; 99US-0145192P.  
PR 23-JUL-1999; 99US-0145145P.



PR 23-JUL-1999; 99US-0145218P.  
 PR 23-JUL-1999; 99US-0145224P.  
 PR 26-JUL-1999; 99US-0145276P.  
 PR 27-JUL-1999; 99US-0145291P.  
 PR 27-JUL-1999; 99US-0145291P.  
 PR 27-JUL-1999; 99US-0145291P.  
 PR 28-JUL-1999; 99US-0145291P.  
 PR 02-AUG-1999; 99US-0146386P.  
 PR 02-AUG-1999; 99US-0146386P.  
 PR 03-AUG-1999; 99US-0147038P.  
 PR 04-AUG-1999; 99US-0147204P.  
 PR 04-AUG-1999; 99US-0147302P.  
 PR 05-AUG-1999; 99US-0147192P.  
 PR 05-AUG-1999; 99US-0147260P.  
 PR 06-AUG-1999; 99US-0147303P.  
 PR 06-AUG-1999; 99US-0147416P.  
 PR 09-AUG-1999; 99US-0147935P.  
 PR 10-AUG-1999; 99US-0148171P.  
 PR 11-AUG-1999; 99US-0148319P.  
 PR 12-AUG-1999; 99US-0148341P.  
 PR 13-AUG-1999; 99US-0148565P.  
 PR 13-AUG-1999; 99US-0148684P.  
 PR 16-AUG-1999; 99US-0149368P.  
 PR 17-AUG-1999; 99US-0149175P.  
 PR 18-AUG-1999; 99US-0149426P.  
 PR 20-AUG-1999; 99US-0149722P.  
 PR 20-AUG-1999; 99US-0149723P.  
 PR 20-AUG-1999; 99US-0149929P.  
 PR 23-AUG-1999; 99US-0149902P.  
 PR 23-AUG-1999; 99US-0149930P.  
 PR 25-AUG-1999; 99US-0150566P.  
 PR 26-AUG-1999; 99US-0150884P.  
 PR 27-AUG-1999; 99US-0151065P.  
 PR 27-AUG-1999; 99US-0151066P.  
 PR 27-AUG-1999; 99US-0151080P.  
 PR 30-AUG-1999; 99US-0151303P.  
 PR 31-AUG-1999; 99US-0151438P.  
 PR 01-SEP-1999; 99US-0151930P.  
 PR 07-SEP-1999; 99US-0152362P.  
 PR 10-SEP-1999; 99US-0153070P.  
 PR 13-SEP-1999; 99US-0153758P.  
 PR 15-SEP-1999; 99US-0154018P.  
 PR 16-SEP-1999; 99US-0154039P.  
 PR 20-SEP-1999; 99US-0154779P.  
 PR 22-SEP-1999; 99US-0155139P.  
 PR 23-SEP-1999; 99US-0155486P.  
 PR 24-SEP-1999; 99US-0155659P.  
 PR 28-SEP-1999; 99US-0156458P.  
 PR 29-SEP-1999; 99US-0156596P.  
 PR 04-OCT-1999; 99US-0157117P.  
 PR 05-OCT-1999; 99US-0157753P.  
 PR 06-OCT-1999; 99US-0157865P.  
 PR 07-OCT-1999; 99US-0158052P.  
 PR 08-OCT-1999; 99US-0158232P.  
 PR 12-OCT-1999; 99US-0158369P.  
 PR 13-OCT-1999; 99US-0159293P.  
 PR 13-OCT-1999; 99US-0159294P.  
 PR 13-OCT-1999; 99US-0159295P.  
 PR 14-OCT-1999; 99US-0159320P.  
 PR 14-OCT-1999; 99US-0159330P.  
 PR 14-OCT-1999; 99US-0159331P.  
 PR 14-OCT-1999; 99US-0159637P.  
 PR 18-OCT-1999; 99US-0159584P.  
 PR 21-OCT-1999; 99US-0160741P.  
 PR 21-OCT-1999; 99US-0160768P.  
 PR 21-OCT-1999; 99US-0160770P.  
 PR 21-OCT-1999; 99US-0160814P.  
 PR 21-OCT-1999; 99US-0160815P.  
 PR 22-OCT-1999; 99US-0160980P.

PR 22-OCT-1999; 99US-0160981P.  
 PR 22-OCT-1999; 99US-0160989P.  
 PR 25-OCT-1999; 99US-0161404P.  
 PR 25-OCT-1999; 99US-0161405P.  
 PR 25-OCT-1999; 99US-0161406P.  
 PR 26-OCT-1999; 99US-0161359P.  
 PR 26-OCT-1999; 99US-0161360P.  
 PR 26-OCT-1999; 99US-0161361P.  
 PR 28-OCT-1999; 99US-0161920P.  
 PR 28-OCT-1999; 99US-0161992P.  
 PR 28-OCT-1999; 99US-0161993P.  
 PR 29-OCT-1999; 99US-0162142P.

Query Match 34.5%; Score 943; DB 3; Length 518;  
 Best Local Similarity 37.5%; Pred. No. 4,4e-83;  
 Matches 196; Conservative 107; Mismatches 195; Indels 24; Gaps 8;

QY 10 IPKVLVSLVLSIVV---KQSLMWREPRKIEHFRSKQGRGPPYHFFIGNVKELVGM 65  
 DB 4 ISTNLTIVLLEFVWSKIKACMILLRLPMLSKFKQGISGPKYKILVGNLSIKKM 63  
 QY 66 MLKSSHPM-PFSNITLPRVLSFYHMKRYGATFLVWPGFTRLVADDLRELF-SK 123  
 DB 64 KKEADLCVLDPNNSNDIPRVFPQYHQMVGDTFLFWGTGKPTIYISNHELAKQVLSK 123  
 QY 124 SEFEKNEAHPLVQLEBDGLSLKGEKMAHKKIISPFHMENTKLIVPVVLYKSVTDNV 183  
 DB 124 FGFITIVKRPVEVILGKGLSFTQGDWIRHRILNPAFSMDRLKAMQPMODCTLRIF 183  
 QY 184 DKWSDKLSNGEV--EVDVYEWFOITEDVYSRTAFSSYEDGRAVFRIOAQOMLCAEA 241  
 DB 184 EEMR-KORNGEVILKIEISKRFKLTADIIATTAAGSYAEGIEICRSQTELEKYIIS 242  
 QY 242 FQKVFIPGVRFPFRGKLRKLDKEIRKSLKLIERRRONALDGGSECKEPA-AKDL 300  
 DB 243 LTNVFIPTGYLPFTNKLKMEHKKVKSIRKILDSRLKS-----KCKTYGVDLL 295  
 QY 301 GLMTQA-----KNVTVDIVIECKSFPPAQKQTSNLLTWTIILSNHEWQAKARDE 353  
 DB 296 GMLTLAKSNVEYKEMDEIIIECKNFYAGGQTSILLTWTIILSLHQWQKLRSE 355  
 QY 354 VLRVGSRDVPTKHVYKLTLSMLNLSKALPPIVATIRRAKSDVKUGYVPCGTE 413  
 DB 356 VFNBCGDKLPDITFESKLTMMNVLMESIRLGPYKISREKTDQMKVGHLEIPKGTGI 415  
 QY 414 LPIIYAVHDOALWGNVNEFPNPARFADGVPRAKHPVGFIPGLGVRTICIGNTALIOA 473  
 DB 416 IIPILKXHRDKALWGEAEQFNPLRPNNGISQATIHNNLLPSPISGPACIAGNFAMVEA 475  
 QY 474 KLTIAVMIGRFTFLAFTYQHAFLVLMVLPQGAIPTRRL 515  
 DB 476 KTVLMTLQQFQSLSPBYKHTPYDHFDPQGLPVMHPL 517

RESULT 6  
 ADA48312  
 ID ADA48312 standard; protein; 454 AA.  
 XX ADA48312;  
 AC 20-NOV-2003 (first entry)  
 XX  
 DT  
 XX  
 DE Rice protein conferring disease resistance in plants.  
 XX disease resistance; pathogen tolerance; plant pathogen; plant; rice.  
 KW  
 XX Oryza sativa.  
 OS  
 XX  
 PN WO200300906-A2.  
 XX  
 PD 03-JAN-2003.  
 XX  
 PF 21-JUN-2002; 2002W0-IB002453.

Page 6

DT	20-NOV-2003 (first entry)
XX	
DE	Rice protein conferring disease resistance in plants.
XX	
KM	disease resistance; pathogen tolerance; plant pathogen; plant; rice.
XX	
OS	<i>Oryza sativa</i> .
XX	
PN	MO2003000906-A2.
XX	
PD	03-JAN-2003.
XX	
FF	21-JUN-2002; 2002WO-1B002453.
XX	
PR	22-JUN-2001; 2001US-0300112P.
XX	
PR	26-SEP-2001; 2001US-0352277P.
PR	22-MAR-2002; 2002US-0366535P.
XX	
2A	(SYGN ) SYNGENTA PARTICIPATIONS AG.
XX	
PI	Glaizebrook J., Briggs S., Cooper B., Goff SA, Moughamer T;
PI	Katagiri F., Kreps J., Provart N., Ricke D., Zhu T;
XX	
DR	WPI: 2003-184052/18.
DR	N-PSDB; ADA48675.
XX	
PT	New polynucleotide comprising a plant nucleotide sequence having an open
PT	reading frame that encodes a polypeptide associated with disease
PT	resistance, useful for conferring resistance or tolerance to a plant
PT	pathogen.

XX	Claim 10; SEQ ID NO 746; 299pp; English.				
XX					
CC	The invention relates to a novel isolated polynucleotide comprising a				
CC	plant nucleotide sequence having an open reading frame that encodes a				
CC	polypeptide associated with disease resistance or its fragment having				
CC	substantially the same activity as the full-length polypeptide. The				
CC	polynucleotide of the invention is useful for conferring resistance or				
CC	tolerance to a plant pathogen. The present sequence represents a protein				
CC	conferring disease resistance used in the invention.				
SO	Sequence 458 AA;				
Query Match	33.7%; Score 919.5; DB 6; Length 458;				
Best Local Similarity	39.5%; Pzed. No. 7,4e-81;				
Matches 188; Conservative	90; Mismatches 153; Indels 43; Gaps 8				
0Y	50 PYPHFFIGVWKEI-----VGMILKASHSMPPSHNILEPVLSPFYHMRKRYGATFLW 102				
DB	14 PIVQIFSGNIGELIKRRDGGAGVLTNLS-----SHDPLFIVQHPFKMILPYGLRFLYW 67				
0Y	103 FGGTFRLITADPDLIRFIS-KSEFYENGEAHLVKQLEGDGILSLGKMAHRRKTIIP 161				
DB	68 FGAQPNICLADVSMWQVLSDRTGYTPKQLTNPHFRLTGLGVLITDODEKRRKRYHP 127				
0Y	162 TPAHKNILKILVPLVLSKTVLDVDDKMSDKL-SENGEVEVDVYYEMFOILTEDVISRTAFGSS 220				
DB	128 AFPMNDLTKVMTMSPDCSRSMWSESELEAKGGLVETLSRRFEELTADVISHTAGSS 187				
0Y	221 YEDGRAVFLQAQOMLLCAEAQKFTFGYAFPPFRGNLKSRLDKELRSILKLIRRR 280				
DB	188 YKGGKQVFLAQRLOFLAFSTFLVQIDGFSTLPTMKRKFKWSLDKRGKMDIITRH 247				
0Y	281 QNA-IDGEGECKEPAAKDLGIMTQANVTVODIVECKSPFAGKQTTSNLTITWTIL 339				
DB	248 ANDQVAGY-----NDILGIMLEAHH-----DTTSHLLTWTMFL 281				

QY 340 LSHMPEHQAAKREIVAKVGSKDVPFKHVAFLKLSLIMBSLNLLELVANRKNED :  
DB 282 LSHPMQEKRLREEIAMECGDK-VPTGMLINKLKVMNFFLEETLRLLSPVALIRRKVDTD 340  
QY 400 VKLGXYIKPGCEELLPIITAVHDQAINGDNVENPRAPDGVPRAAKHVGFIIPPLG 459  
::|||:|||:|||::|||:|||:|||:|||||:

DB 341 IEIIGIMPEGALITPIATIRHDKVEGEDADEPFRPEFENGVTBAAGHPALLSFSG 400  
OY 460 VRTICGNAILLOAKLTIAVMIOFTPHLAFTYQAPVTMLLYPHGAPITFRLL 515  
DB 401 PRSCICGNFAMIEKAVIAMILORFSPILSPKYVHAFTVITLPRYGLPMILKSL 456

RESULT 8  
AAG31393  
ID AAG31393 standard; protein: 484 AA.  
XX AAG31393;  
AC AAG31393;  
DT 17-OCT-2000 (first entry)  
XX  
DE Arabidopsis thaliana protein fragment SEQ ID NO: 37692.  
XX  
KW Protein identification; signal transduction pathway; metabolic pathway;  
KM hybridisation assay; genetic mapping; gene expression control; promoter;  
XX termination sequence.  
OS Arabidopsis thaliana.  
XX EP1033405-A2.  
XX  
PD 06-SEP-2000.  
XX  
PF 25-FEB-2000; 2000EP-00301439.  
XX  
XX 25-FEB-1999; 99US-011825P.  
PR 05-MAR-1999; 99US-0123180P.  
PR 09-MAR-1999; 99US-0123548P.  
PR 23-MAR-1999; 99US-0125788P.  
PR 25-MAR-1999; 99US-0126264P.  
PR 29-MAR-1999; 99US-0126785P.  
PR 01-APR-1999; 99US-0127462P.  
PR 06-APR-1999; 99US-0128234P.  
PR 08-APR-1999; 99US-0128714P.  
PR 16-APR-1999; 99US-0129845P.  
PR 19-APR-1999; 99US-0130077P.  
PR 21-APR-1999; 99US-0130449P.  
PR 23-APR-1999; 99US-0130510P.  
PR 23-APR-1999; 99US-0130891P.  
PR 28-APR-1999; 99US-0131448P.  
PR 30-APR-1999; 99US-0132048P.  
PR 04-MAY-1999; 99US-0132407P.  
PR 05-MAY-1999; 99US-0132484P.  
PR 06-MAY-1999; 99US-0132486P.  
PR 06-MAY-1999; 99US-0132487P.  
PR 07-MAY-1999; 99US-0132863P.  
PR 11-MAY-1999; 99US-0132866P.  
PR 14-MAY-1999; 99US-0134218P.  
PR 14-MAY-1999; 99US-0134219P.  
PR 14-MAY-1999; 99US-0134221P.  
PR 14-MAY-1999; 99US-0134370P.  
PR 18-MAY-1999; 99US-0134768P.  
PR 19-MAY-1999; 99US-0135941P.  
PR 20-MAY-1999; 99US-0135124P.  
PR 21-MAY-1999; 99US-0135353P.  
PR 24-MAY-1999; 99US-0135629P.  
PR 25-MAY-1999; 99US-0136021P.  
PR 27-MAY-1999; 99US-0136392P.  
PR 28-MAY-1999; 99US-0136783P.  
PR 01-JUN-1999; 99US-0137223P.  
PR 03-JUN-1999; 99US-0137528P.  
PR 04-JUN-1999; 99US-0137502P.  
PR 07-JUN-1999; 99US-0137724P.  
PR 08-JUN-1999; 99US-0138094P.  
PR 10-JUN-1999; 99US-0138540P.  
PR 10-JUN-1999; 99US-0138647P.  
PR 14-JUN-1999; 99US-0139119P.  
PR 16-JUN-1999; 99US-0139452P.

PR 16-JUN-1999; 99US-0139453P.  
PR 17-JUN-1999; 99US-0139492P.  
PR 18-JUN-1999; 99US-0139454P.  
PR 18-JUN-1999; 99US-0139455P.  
PR 18-JUN-1999; 99US-0139456P.  
PR 18-JUN-1999; 99US-0139457P.  
PR 18-JUN-1999; 99US-0139458P.  
PR 18-JUN-1999; 99US-0139459P.  
PR 18-JUN-1999; 99US-0139460P.  
PR 18-JUN-1999; 99US-0139461P.  
PR 18-JUN-1999; 99US-0139462P.  
PR 18-JUN-1999; 99US-0139463P.  
PR 18-JUN-1999; 99US-0139750P.  
PR 18-JUN-1999; 99US-0139763P.  
PR 21-JUN-1999; 99US-0139817P.  
PR 22-JUN-1999; 99US-0139899P.  
PR 23-JUN-1999; 99US-0140353P.  
PR 23-JUN-1999; 99US-0140354P.  
PR 24-JUN-1999; 99US-0140695P.  
PR 28-JUN-1999; 99US-0140823P.  
PR 29-JUN-1999; 99US-0140991P.  
PR 30-JUN-1999; 99US-0141287P.  
PR 01-JUL-1999; 99US-0141842P.  
PR 01-JUL-1999; 99US-0142154P.  
PR 02-JUL-1999; 99US-0142055P.  
PR 06-JUL-1999; 99US-0142390P.  
PR 08-JUL-1999; 99US-0142803P.  
PR 09-JUL-1999; 99US-0142920P.  
PR 12-JUL-1999; 99US-0142977P.  
PR 13-JUL-1999; 99US-0143542P.  
PR 14-JUL-1999; 99US-0143624P.  
PR 15-JUL-1999; 99US-0144005P.  
PR 16-JUL-1999; 99US-0144085P.  
PR 16-JUL-1999; 99US-0144086P.  
PR 19-JUL-1999; 99US-0144325P.  
PR 19-JUL-1999; 99US-0144331P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 19-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144332P.  
PR 20-JUL-1999; 99US-0144632P.  
PR 21-JUL-1999; 99US-0144814P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 21-JUL-1999; 99US-0145088P.  
PR 22-JUL-1999; 99US-0145085P.  
PR 22-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145089P.  
PR 22-JUL-1999; 99US-0145192P.  
PR 23-JUL-1999; 99US-0145145P.  
PR 23-JUL-1999; 99US-0145218P.  
PR 23-JUL-1999; 99US-0145276P.  
PR 26-JUL-1999; 99US-0145913P.  
PR 27-JUL-1999; 99US-0145918P.  
PR 27-JUL-1999; 99US-0145951P.  
PR 28-JUL-1999; 99US-0145951P.  
PR 02-AUG-1999; 99US-0146386P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 02-AUG-1999; 99US-0146389P.  
PR 03-AUG-1999; 99US-0147038P.  
PR 04-AUG-1999; 99US-0147204P.  
PR 04-AUG-1999; 99US-0147130P.  
PR 05-AUG-1999; 99US-0147132P.  
PR 05-AUG-1999; 99US-0147260P.  
PR 06-AUG-1999; 99US-0147303P.  
PR 06-AUG-1999; 99US-0147416P.  
PR 09-AUG-1999; 99US-0147453P.  
PR 09-AUG-1999; 99US-0147935P.  
PR 10-AUG-1999; 99US-0148171P.  
PR 11-AUG-1999; 99US-0148319P.  
PR 12-AUG-1999; 99US-0148341P.

PR 13-AUG-1999; 99US-0148565P.  
PR 13-AUG-1999; 99US-0148684P.  
PR 16-AUG-1999; 99US-0149368P.  
PR 17-AUG-1999; 99US-0149175P.  
PR 18-AUG-1999; 99US-0149426P.  
PR 20-AUG-1999; 99US-0149722P.  
PR 20-AUG-1999; 99US-0149723P.  
PR 20-AUG-1999; 99US-0149922P.  
PR 23-AUG-1999; 99US-0149902P.  
PR 23-AUG-1999; 99US-0149930P.  
PR 25-AUG-1999; 99US-0150566P.  
PR 26-AUG-1999; 99US-0150884P.  
PR 27-AUG-1999; 99US-0151065P.  
PR 27-AUG-1999; 99US-0151066P.  
PR 27-AUG-1999; 99US-0151080P.  
PR 30-AUG-1999; 99US-0151303P.  
PR 31-AUG-1999; 99US-0151438P.  
PR 01-SEP-1999; 99US-0151930P.  
PR 07-SEP-1999; 99US-0152363P.  
PR 10-SEP-1999; 99US-0153070P.  
PR 13-SEP-1999; 99US-0153758P.  
PR 15-SEP-1999; 99US-0154018P.  
PR 16-SEP-1999; 99US-0154039P.  
PR 20-SEP-1999; 99US-0154779P.  
PR 22-SEP-1999; 99US-0155139P.  
PR 23-SEP-1999; 99US-0155486P.  
PR 24-SEP-1999; 99US-0156559P.  
PR 28-SEP-1999; 99US-0156588P.  
PR 29-SEP-1999; 99US-0156596P.  
PR 04-OCT-1999; 99US-0157117P.  
PR 05-OCT-1999; 99US-0157753P.  
PR 06-OCT-1999; 99US-0157865P.  
PR 07-OCT-1999; 99US-0158029P.  
PR 08-OCT-1999; 99US-0158232P.  
PR 12-OCT-1999; 99US-0158369P.  
PR 13-OCT-1999; 99US-0159293P.  
PR 13-OCT-1999; 99US-0159294P.  
PR 13-OCT-1999; 99US-0159295P.  
PR 14-OCT-1999; 99US-0159329P.  
PR 14-OCT-1999; 99US-0159330P.  
PR 14-OCT-1999; 99US-0159331P.  
PR 14-OCT-1999; 99US-0159637P.  
PR 14-OCT-1999; 99US-0159638P.  
PR 18-OCT-1999; 99US-0159584P.  
PR 21-OCT-1999; 99US-0160741P.  
PR 21-OCT-1999; 99US-0160767P.  
PR 21-OCT-1999; 99US-0160768P.  
PR 21-OCT-1999; 99US-0160770P.  
PR 21-OCT-1999; 99US-0160814P.  
PR 21-OCT-1999; 99US-0160815P.  
PR 22-OCT-1999; 99US-0160800P.  
PR 22-OCT-1999; 99US-0160981P.  
PR 22-OCT-1999; 99US-0160989P.  
PR 25-OCT-1999; 99US-0161404P.  
PR 25-OCT-1999; 99US-0161405P.  
PR 25-OCT-1999; 99US-0161406P.  
PR 26-OCT-1999; 99US-0161359P.  
PR 26-OCT-1999; 99US-0161360P.  
PR 26-OCT-1999; 99US-0161361P.  
PR 28-OCT-1999; 99US-0161920P.  
PR 28-OCT-1999; 99US-0161922P.  
PR 28-OCT-1999; 99US-0161993P.  
PR 29-OCT-1999; 99US-0162142P.

Query Match 33.6%; Score 917; DB 3; Length 484;  
Best Local Similarity 38.2%; Pred. No. 1,4e-80;  
Matches 187; Conservative 100; Mismatches 183; Indels 20; Gaps 7;

QY 38 IEHFSKQIGRPFYHFFIGNVKELVGMMLKASSHPM-PFSHNILPRLVLSFYHWRKJYG 96  
DB 2 LSKRPFKQGISGPKYKILYGNLSFKKKKKKALCVLDPNSNDIFPRVFPQYHQMWSQYG 61  
QY 97 ATFLVWGFTRFLTVADPDILIRELF-SKSEFYKNAHPLVYQLEDDGLSLKGGKMAHH 155

DB 62 DTFLFWGTGKPTLYISNHELAKOVLSKSGFTIIPVAREBEVFLIFGKGSLFGOGDMWRH 121  
QY 156 KRISPTFHENKLVLPVVLKSVTMDVQKSKLSSENGEV--EVDVVEWFOITFEDVIS 213  
DB 122 RRLNPAFSMDRLKAMQFPGDCTLRIFEEWR-KQRNNGEVLKIKISGFHKLTDILA 180  
QY 214 RTAFSSYEDGRAVFRILQAOOMLCAEAFQKVEIPGVFFPTGNLKSRLDKYIRKSL 273  
DB 181 TTAFGSSAEBIEELCRQTELEKYYISLNVFIPGQYLPPTNLKLWELHKKVNGSLK 240  
QY 274 KLIERRONAIDGGECEKBP-AKDLGLMIA-----KNVTQDIYECKSPFFAG 325  
DB 241 RIIDSRLKS-----KCKTYGYGDDLLGYMLTPAKSNEYERKMRDEIIECKNFYAG 293  
QY 326 KOTSNLLTWTTLTILSNHPEWQAKARDEVLRVCSRDVPTKDHVYKLTLSMILNESLRI 385  
DB 294 OGTSILLTWTMLSLHQGQEKLRREVNECKKIPDTDFSKLKNAMVLMESLNL 353  
QY 386 YPPIVATIRRAKSDVKLGKIKIPGCTELLIPITAVHDDALINGNDVNEFPARFAGVIR 445  
DB 354 YGPVAKISREATQDMKYGHEIIPKGTSLIIPLMKMRDKAIWGEDAEQFPLRFENGISO 413  
QY 446 AAKHVGFIIPRGLVPRGICIGONLAILQAKLTIAVIGRFFPHLAPTYOHAFTVIMLYPO 505  
DB 414 ATIHFNALLPFSIGPRACIAKNFMWAKVTLMILOQFQSLSPFEYKHTPVDFHDLFPQ 473  
QY 506 HGAPITFRRL 515  
DB 474 YGLPVTMLHPL 483

RESULT 9  
AAG31394  
ID AAG31394 standard; protein; 456 AA.  
XX  
AC AAG31394;  
XX  
DT 17-OCT-2000 (first entry)  
XX  
DE Arabidopsis thaliana protein fragment SEQ ID NO: 37693.  
XX  
KW Protein identification; signal transduction pathway; metabolic pathway;  
hybridisation assay; genetic mapping; gene expression control; promoter;  
termination sequence.  
XX  
OS Arabidopsis thaliana.  
XX  
PN EP1033405-A2.  
XX  
PD 06-SEP-2000.  
XX  
PF 25-FEB-2000; 2000EP-00301439.  
XX  
PR 25-FEB-1999; 99US-0121825P.  
PR 05-MAR-1999; 99US-0123180P.  
PR 09-MAR-1999; 99US-0123548P.  
PR 23-MAR-1999; 99US-0125788P.  
PR 25-MAR-1999; 99US-0126264P.  
PR 25-MAR-1999; 99US-0126264P.  
PR 01-APR-1999; 99US-0127462P.  
PR 06-APR-1999; 99US-0128234P.  
PR 08-APR-1999; 99US-0128714P.  
PR 16-APR-1999; 99US-0129845P.  
PR 19-APR-1999; 99US-0130077P.  
PR 21-APR-1999; 99US-0130449P.  
PR 23-APR-1999; 99US-0130510P.  
PR 23-APR-1999; 99US-0130891P.  
PR 28-APR-1999; 99US-0131449P.  
PR 30-APR-1999; 99US-0132048P.  
PR 30-APR-1999; 99US-0132407P.  
PR 04-MAY-1999; 99US-0132484P.  
PR 05-MAY-1999; 99US-0132485P.

Mon May 24 08:10:09 2004

us-09-992-901-2.rag

PR 06-MAY-1999; 99US-0132486P.  
PR 06-MAY-1999; 99US-0132487P.  
PR 07-MAY-1999; 99US-0132863P.  
PR 11-MAY-1999; 99US-0134256P.  
PR 14-MAY-1999; 99US-0134218P.  
PR 14-MAY-1999; 99US-0134421P.  
PR 14-MAY-1999; 99US-0134221P.  
PR 14-MAY-1999; 99US-0134370P.  
PR 18-MAY-1999; 99US-0134768P.  
PR 19-MAY-1999; 99US-0134941P.  
PR 20-MAY-1999; 99US-0135124P.  
PR 21-MAY-1999; 99US-0135353P.  
PR 24-MAY-1999; 99US-0135629P.  
PR 25-MAY-1999; 99US-0136021P.  
PR 27-MAY-1999; 99US-0136392P.  
PR 28-MAY-1999; 99US-0136782P.  
PR 01-JUN-1999; 99US-0137222P.  
PR 03-JUN-1999; 99US-0137528P.  
PR 04-JUN-1999; 99US-0137502P.  
PR 07-JUN-1999; 99US-0137724P.  
PR 08-JUN-1999; 99US-0138094P.  
PR 10-JUN-1999; 99US-0138540P.  
PR 10-JUN-1999; 99US-0138847P.  
PR 14-JUN-1999; 99US-0139119P.  
PR 16-JUN-1999; 99US-0139452P.  
PR 16-JUN-1999; 99US-0139452P.  
PR 17-JUN-1999; 99US-0139482P.  
PR 17-JUN-1999; 99US-0139482P.  
PR 18-JUN-1999; 99US-0139454P.  
PR 18-JUN-1999; 99US-0139455P.  
PR 18-JUN-1999; 99US-0139456P.  
PR 18-JUN-1999; 99US-0139457P.  
PR 18-JUN-1999; 99US-0139458P.  
PR 18-JUN-1999; 99US-0139459P.  
PR 18-JUN-1999; 99US-0139460P.  
PR 18-JUN-1999; 99US-0139461P.  
PR 18-JUN-1999; 99US-0139462P.  
PR 18-JUN-1999; 99US-0139463P.  
PR 18-JUN-1999; 99US-0139763P.  
PR 18-JUN-1999; 99US-0139763P.  
PR 21-JUN-1999; 99US-0139817P.  
PR 22-JUN-1999; 99US-0139859P.  
PR 23-JUN-1999; 99US-0140353P.  
PR 23-JUN-1999; 99US-0140354P.  
PR 24-JUN-1999; 99US-0140695P.  
PR 28-JUN-1999; 99US-0140823P.  
PR 29-JUN-1999; 99US-0140991P.  
PR 30-JUN-1999; 99US-0141287P.  
PR 01-JUL-1999; 99US-0141842P.  
PR 01-JUL-1999; 99US-0142154P.  
PR 02-JUL-1999; 99US-0142055P.  
PR 06-JUL-1999; 99US-0142390P.  
PR 08-JUL-1999; 99US-0142803P.  
PR 09-JUL-1999; 99US-0142920P.  
PR 12-JUL-1999; 99US-0142977P.  
PR 13-JUL-1999; 99US-0143542P.  
PR 14-JUL-1999; 99US-0143624P.  
PR 15-JUL-1999; 99US-0144005P.  
PR 16-JUL-1999; 99US-0144086P.  
PR 16-JUL-1999; 99US-0144086P.  
PR 19-JUL-1999; 99US-0144331P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 20-JUL-1999; 99US-0144352P.  
PR 20-JUL-1999; 99US-0144632P.  
PR 20-JUL-1999; 99US-0144864P.  
PR 20-JUL-1999; 99US-0144814P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 22-JUL-1999; 99US-0145085P.  
PR 22-JUL-1999; 99US-0145087P.

PR 22-JUL-1999; 99US-0145089P.  
PR 22-JUL-1999; 99US-0145192P.  
PR 23-JUL-1999; 99US-0145145P.  
PR 23-JUL-1999; 99US-0145218P.  
PR 23-JUL-1999; 99US-0145224P.  
PR 26-JUL-1999; 99US-0145276P.  
PR 27-JUL-1999; 99US-0145913P.  
PR 27-JUL-1999; 99US-0145918P.  
PR 27-JUL-1999; 99US-0145919P.  
PR 28-JUL-1999; 99US-0145951P.  
PR 02-AUG-1999; 99US-0146386P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 02-AUG-1999; 99US-0146389P.  
PR 03-AUG-1999; 99US-0147038P.  
PR 04-AUG-1999; 99US-0147204P.  
PR 04-AUG-1999; 99US-0147302P.  
PR 05-AUG-1999; 99US-0147192P.  
PR 05-AUG-1999; 99US-0147260P.  
PR 06-AUG-1999; 99US-0147303P.  
PR 06-AUG-1999; 99US-0147416P.  
PR 09-AUG-1999; 99US-0147935P.  
PR 09-AUG-1999; 99US-0148171P.  
PR 10-AUG-1999; 99US-0148319P.  
PR 11-AUG-1999; 99US-0148341P.  
PR 12-AUG-1999; 99US-0148565P.  
PR 13-AUG-1999; 99US-0148684P.  
PR 13-AUG-1999; 99US-0149368P.  
PR 16-AUG-1999; 99US-0149175P.  
PR 17-AUG-1999; 99US-0149426P.  
PR 18-AUG-1999; 99US-0149722P.  
PR 20-AUG-1999; 99US-0149723P.  
PR 20-AUG-1999; 99US-0149929P.  
PR 23-AUG-1999; 99US-0149902P.  
PR 23-AUG-1999; 99US-0149930P.  
PR 25-AUG-1999; 99US-0150066P.  
PR 26-AUG-1999; 99US-0150084P.  
PR 27-AUG-1999; 99US-0151065P.  
PR 27-AUG-1999; 99US-0151066P.  
PR 27-AUG-1999; 99US-0151080P.  
PR 30-AUG-1999; 99US-0151303P.  
PR 31-AUG-1999; 99US-0151438P.  
PR 01-SEP-1999; 99US-0151930P.  
PR 07-SEP-1999; 99US-0152363P.  
PR 10-SEP-1999; 99US-0153070P.  
PR 13-SEP-1999; 99US-0153758P.  
PR 15-SEP-1999; 99US-0154018P.  
PR 16-SEP-1999; 99US-0154039P.  
PR 20-SEP-1999; 99US-0155139P.  
PR 22-SEP-1999; 99US-0155486P.  
PR 23-SEP-1999; 99US-0155659P.  
PR 24-SEP-1999; 99US-0156458P.  
PR 28-SEP-1999; 99US-0156596P.  
PR 29-SEP-1999; 99US-0157117P.  
PR 04-OCT-1999; 99US-0157753P.  
PR 05-OCT-1999; 99US-0157655P.  
PR 06-OCT-1999; 99US-0158029P.  
PR 07-OCT-1999; 99US-0158212P.  
PR 08-OCT-1999; 99US-0158369P.  
PR 12-OCT-1999; 99US-0159293P.  
PR 13-OCT-1999; 99US-0159294P.  
PR 13-OCT-1999; 99US-0159294P.  
PR 13-OCT-1999; 99US-0159295P.  
PR 14-OCT-1999; 99US-0159339P.  
PR 14-OCT-1999; 99US-0159331P.  
PR 14-OCT-1999; 99US-0159637P.  
PR 14-OCT-1999; 99US-0159684P.  
PR 18-OCT-1999; 99US-0159584P.  
PR 21-OCT-1999; 99US-0160741P.  
PR 21-OCT-1999; 99US-0160767P.  
PR 21-OCT-1999; 99US-0160768P.  
PR 21-OCT-1999; 99US-0160770P.

PR 21-OCT-1999; 99US-0160814P.  
 PR 21-OCT-1999; 99US-0160815P.  
 PR 22-OCT-1999; 99US-0160980P.  
 PR 22-OCT-1999; 99US-0160981P.  
 PR 22-OCT-1999; 99US-0160982P.  
 PR 25-OCT-1999; 99US-0161404P.  
 PR 25-OCT-1999; 99US-0161405P.  
 PR 25-OCT-1999; 99US-0161406P.  
 PR 26-OCT-1999; 99US-0161359P.  
 PR 26-OCT-1999; 99US-0161360P.  
 PR 26-OCT-1999; 99US-0161361P.  
 PR 28-OCT-1999; 99US-0161920P.  
 PR 28-OCT-1999; 99US-0161922P.  
 PR 28-OCT-1999; 99US-0161932P.  
 PR 29-OCT-1999; 99US-0162142P.

Query Match 31.7%; Score 864.5; DB 3; Length 456;  
 Best Local Similarity 38.5%; Pred. No. 1,88-75;  
 Matches 174; Conservative 94; Mismatches 165; Indels 19; Gaps 6;

QY 75 PESHNLPRLVSFYHWRKTYGATFLVFGPTFLTVADPDLIREIF-SKSEFYKKNKAH 133  
 DB 12 PMSNDIFRPFVPOYHOMNSQYCDTFLFWTGKPTIYISNHELAKQVLSKKGFTIIPVKR 71  
 QY 134 PLVKQLESGDGLSLGKGAHHRKTIPTFMENIKLVVVLKSVTDMDKMDKSEN 193  
 DB 72 PEVFTILFGKGLSPFGDDIRRRRLINPAFSMDRLKAMTOPMGDCTLRIFEWR-KORRN 130  
 QY 194 GEV--EVDYVEFQILTEDVISRATFAGSSYEDGSAVFLQAOQMLCAEAFQKFIPIGR 251  
 DB 131 GEVLKIRSKFHLTDITATFAGSSYAGELICSCQLEKVIYISLTNFTIGTQ 130  
 QY 252 PPTFRGNLSKRLDKIRKSLIKLERRRQNALIDGEGCEPA-ADULGLMLQA--- 306  
 DB 191 YLPTPTNLKWEIHKVKNISIRIIDSRLKS-----KCTYGYGDLGLGMLTAKSN 243  
 QY 307 ---KNTVGDIVIECKSPFAGKQTSNMLTWTTLISMHQMOAKADEVLARCGRDV 363  
 DB 244 EYERKRMDEIIECCNFYFAGQTSILITWTTLISLHGMQEKLEBVFNCGDOKI 303  
 QY 364 PTKDHVVKLTSLMILNESLRYPPIVATIRRAKSDVKGKIPCGTELLPIAVHHD 423  
 DB 304 PDDTFSKLTMMVMYMSLRLYGPVIXISREATQDMKVLBRPKGTSLIIPILKQHRD 363  
 QY 424 QATWMDVNEFNPAPRADGVPRAAHPVCGFIFPGVGTCTIGONLALIOAKLTAVIQR 483  
 DB 364 KALWGEDAQENFLREFNGISQATLHPNALPFSISGPRACIAKQFAMWEAKTVLMILOQ 423  
 QY 484 FTEHLAPTYQHAFTVLMMLYPOHGAFTTFRRL 515  
 DB 424 FQLSLSPRYKHTPVHFDLPQYGLFVMLHPL 455

RESULT 10  
 AB05536  
 ID AB05536 standard; protein; 505 AA.  
 XX AB05536;  
 XX  
 DT 14-NOV-2002 (first entry)  
 XX  
 DE Breast cancer-associated protein 1.  
 XX  
 KM Breast cancer; breast cancer-associated gene sequence; drug development;  
 KM pharmacogenetics; biosensor development.  
 XX  
 OS Unidentified.  
 XX  
 XX W0200259377-A2.  
 XX  
 XX 01-AUG-2002.  
 XX  
 XX  
 PF 24-JAN-2002; 2002W0-US002242.

XX 24-JAN-2001; 2001US-0263965P.  
 PR 02-FEB-2001; 2001US-0265928P.  
 PR 09-APR-2001; 2001US-00829472.  
 PR 09-APR-2001; 2001US-0282658P.  
 PR 04-MAY-2001; 2001US-0288590P.  
 PR 29-MAY-2001; 2001US-0294443P.  
 XX  
 PA (EOSB-) EOS BIOTECHNOLOGY INC.  
 XX  
 PI Mack DH, Glen KC, Afar D;  
 XX  
 DR WPI, 2002-583738/62.  
 DR N-PSDB; AB07693.  
 PT  
 PT Detecting a breast cancer-associated transcript in a patient's cell,  
 PT useful for diagnosing breast cancer, comprises contacting a biological  
 PT sample with a polynucleotide that selectively hybridizes with breast  
 PT cancer nucleic acids.

Dislosure; Page 347; 414pp; English.

The invention comprises a method of detecting a breast cancer-associated transcript in a cell from a patient. The method of the invention involves contacting a biological sample from the patient with a nucleotide that hybridizes to one of the 69 breast cancer-associated gene sequences shown in the specification. The method of the invention is useful in the diagnosis or prognosis of breast cancer, and for detecting genes that are up or down-regulated in breast cancer cells. Genes identified by the method of the invention can be used in diagnostic purposes and also as targets for screening for therapeutic compounds that modulate breast cancer (e.g. hormones or antibodies). Identification of genes that are over or under expressed in breast cancer can additionally provide high-resolution, high-sensitivity datasets which can be used in the areas of diagnostics, therapeutics, drug development, pharmacogenetics, protein structure and biosensor development. Amino acid sequences AB05536 - AB05604 represent the proteins encoded by the 69 breast cancer-associated genes of the invention

Sequence 505 AA;

Query Match 19.9%; Score 543.5; DB 5; Length 505;  
 Best Local Similarity 29.3%; Pred. No. 6,18-44;  
 Matches 149; Conservative 104; Mismatches 155; Indels 61; Gaps 20;

QY 11 PAVLIVSLVLYKWSMLM--WRPKIEHFSKQGRGPPYHFFIGNVXELVGMML 67  
 DB 13 PVLILLCMSLILFQVIRLYQRRRMIRAL--HL-----PPAPAHVFYGH-KEF----- 60  
 QY 68 KASHPMFPSSHNLPRVLSFYHWRKTYGATFLVWRBP-TPRLTVADPDLIREIFSKSEF 126  
 DB 61 ----YPV-----KEPEVYHKLMEXYPCAVPLMVGPFTMFSPVHPDPAK-ILLKQD 107  
 QY 127 YEKNAHPLVVKOLEGGLSLGKGAHHRKTIPTFMENIKLVVVLKSVTDMVQX 186  
 DB 108 PKSAVSHKILBSWVRGLVTLDSGKMKKHQIVAPGNISILKFTIMSESVYMLNWK 167  
 QY 187 SDKLSNGEVEVDYEMFQILTEDVISRATF--GSSYEDG-----RAYFRQA---QC 234  
 DB 168 EERIAQNSRLK--LFQVSLMTLDSIMKCAFSHQSGS:QLDSTDSYKAAVFNLSKISNR 225  
 QY 235 MLCALAEQKVFIPGRFPFRGNLSKRLDKIRKSLIKLERRRQNALIDGEGCEKPEP 294  
 DB 226 MNPLFHHVDYF-----KFSOGQILFS-KENQELHQFTENKIQDRKESLKDIXQDPTQK 279  
 QY 295 AAKDLGLMIOAKNTVODIVE-----ECSFFPAGKQTSNMLTWTTLISMPPEWQK 349  
 DB 280 RRMDFDLITLSAKSENTDFSEADLOAEVKTFMFGHDTTSASISWLVCYLAQYBEHQOR 339  
 QY 350 ARDEVLYRCGRSDVPTKHVVKLTSLMILNESLRYPPIVATIRRAKSDVKL-GGYKIP 408  
 DB 340 CRDEIRELLDGGSSITWHLSQMPYTTMCICGRILVAFVNVISRLDLKDPITFPDGSGLP 399

QY 409 CGTELLIPILAVHDDQAIWGVNVEFNFARFADGVPPAAK-HPVGFIPEGLGYRTICGN 467  
DB 400 AGITVFIMWALHNHPYFW-EDPOVFNPLFRSR--ENSEKHHPYAFIPFSGJRNICIGH 456  
QY 468 LALIOAKTLTAVMIORFTEHAPYQAP 496  
DB 457 FALIECKVAVALTLIR--FKLAPDSRPP 483

RESULT 11  
AAE14447  
ID AAE14447 standard; protein; 505 AA.  
XX AAE14447;  
AC  
XX  
XX 26-MAR-2002 (first entry)  
DE Human drug metabolizing enzyme (DME)-10.  
XX  
XX Human; drug metabolizing enzyme; DME-10; autoimmune; inflammatory;  
XX cell proliferative; developmental; endocrine; eye; metabolic; AIDS;  
XX gastrointestinal disorder; liver disorder; cancer; arteriosclerosis;  
XX adult respiratory distress syndrome; anaemia; epilepsy; hypothyroidism;  
XX hypochlasmia; pituitary; diabetes; hypogonadism; conjunctivitis;  
XX glaucoma; cystic fibrosis; hypercholesterolaemia; gastritis;  
XX peptic ulcer; hepatitis; gene therapy.  
XX  
XX Homo sapiens.  
OS  
XX  
XX Key Location/Qualifiers  
FH 1..31  
FT /label=Signal\_peptide  
FT 14..33  
FT /label=Transmembrane\_domain  
FT 32..505  
FT /label=Mature\_DME-10  
PN WO200190334-A2.  
XX  
XX 29-NOV-2001.  
PD  
XX  
XX 25-MAY-2001; 2001WO-US017150.  
XX  
XX 25-MAY-2000; 2000US-0207901P.  
XX  
XX 01-JUN-2000; 2000US-0208983P.  
XX  
XX 07-JUN-2000; 2000US-0209861P.  
XX  
XX 15-JUN-2000; 2000US-0211825P.  
XX  
XX 22-JUN-2000; 2000US-0213744P.  
XX  
XX (INCY-) INCYTE GENOMICS INC.  
PA  
XX  
XX Yue H, Sanjanwala WS, Baughn MR, Gandhi AR, Ring HZ, Elliott V;  
XX  
XX Maia NK, Yang J, Khan PA, Ramkumar T, Tang YT, Hafalla A, Lal P;  
XX  
XX Nguyen DB, Vao MG, Lee EA, Tribouley CM, Patterson C, Lu Y,  
XX  
XX Burford N, Ding L, Bruns CM, Kearney L, Reddy R;  
XX  
XX WPI; 2002-097650/13.  
XX  
XX N-PSDB; AAD24015.  
XX  
XX New human drug metabolizing enzymes and polynucleotides encoding the  
XX  
XX enzyme for diagnosing, preventing or treating cell proliferative,  
XX  
XX autoimmune/inflammatory, endocrine, eye, metabolic and gastrointestinal  
XX  
XX disorders.  
XX  
XX  
XX Claim 1; Page 145-146; 158pp; English.

CC autoimmune hemolytic anaemia, autoimmune thyroiditis, Crohn's disease,  
CC atopic dermatitis, diabetic mellitus, Graves' disease, systemic lupus  
CC glomerulonephritis, rheumatoid arthritis, scleroderma, haemodialysis and  
CC erythematous, systemic sclerosis, ulcerative colitis, haemodialysis and  
CC uveitis, viral, bacterial, fungal, parasitic, protozoal, helminthic  
CC infections and trauma, and cell proliferative disorders such as cancer,  
CC actinic keratosis, arteriosclerosis, atherosclerosis, bursitis,  
CC cirrhosis, hepatitis and psoriasis. Developmental disorders include  
CC anaemia, renal tubular acidosis, epilepsy, hypothyroidism and cataract,  
CC and endocrine disorders include disorders of hypothalamus and pituitary,  
CC disorders associated with hypopituitarism, including sarcoidosis,  
CC diabetes insipidus, hypogonadism, disorders associated with  
CC hypothyroidism including goitre, acute thyroiditis, Graves' disease,  
CC disorders associated with hyperparathyroidism, pancreatic disorders such  
CC as type I or type II diabetes mellitus, disorders associated with  
CC adrenals such as hyperplasia, Cushing's disease, endometriosis,  
CC infertility, hypergonadal disorders, and gynaecomastia. Eye disorders  
CC include conjunctivitis, keratitis, glaucoma and macular degeneration, and  
CC metabolic disorders include diabetes, cystic fibrosis, goitre,  
CC hypercholesterolaemia, hypoglycaemia, hypercalcaemia, lysosomal storage  
CC diseases, obesity, phenylketonuria and hypocalcaemia. Also the molecules  
CC are useful for treating gastrointestinal disorders such as dysphagia,  
CC gastritis, peptic ulcer, cholelithiasis, cirrhosis, hepatitis,  
CC hyperbilirubinaemia, constipation, diarrhoea, jaundice, Wilson's disease,  
CC thrombosis and hepatic tumours. The DME polypeptide is also useful for  
CC screening its agonist or antagonist  
XX  
XX Sequence 505 AA;  
SQ

Query Match 19.9%; Score 543.5; DB 5; Length 505;  
Best Local Similarity 29.3%; Pred. No. 6.1e-44;  
Matches 149; Conservative 104; Mismatches 195; Indels 61; Gaps 20;

QY 11 PKVLVSLVSLVVKMSLLM--WRPRIEBHFSQGRGPPYFPIGNVETLVGMML 67  
DB 13 PFLITLILCWSLPLFQVIRLYKORRRWIRPL--HL-----FPAPPAWFTYH--KEF----- 60  
QY 68 KASHMPMPFSNHLPRVLSFYHNRKLYGATFLWFGP-TFLTVADPDILIREIFSKSEF 126  
DB 61 ----YPV-----KEFEVYHKLMEXPCAVPLWGPFTMFSVHDPDAK--ILKRPD 107  
QY 127 YEKNEAHLVYKLEGDGLSLKGEKMAHKKIISPTFHEMNLKLAVPVLYKSTLDMWDK 186  
DB 108 PKSAVSHKTESWVGRLVTLDSGKMKKHQIVKPGNISILKIFITMESRYRMNLTK 167  
QY 187 SDKLSNGEVEVDYEWFOILTEDVLSRTAF---GSGSYEDG-----RAVFLQA---QQ 234  
DB 168 EERLAQNGRL--LFQVSLMTLDSIMKCAFSGHQSIOIDSTLDSYKAVFNLSKISNGR 225  
QY 235 MLCAPAKQYFIFGAYFPFRGLTKSRKLDKIRKSLTLIRRRONALDGESECKE 294  
DB 226 MNFLHNDLVF---KFSQGIQS-KRNOEHQTEKVTODRKSLKDKUKOTTK 279  
QY 295 AAKDLIGMIQAKVTVODIVE-----ECKSFPAKQKTSNLLTWTTLLSNHPSWQAK 349  
DB 280 RRMDELILLAKENKRDSEADLQAEVTFMFGHDTSSAISWILYCLAKYPHQQR 339  
QY 350 ARDEVLRVGSRDVPTDHYVYKLTLSNINLESRLYPPVATIRAKSVYL--GGYKIP 408  
DB 340 CRDEIRELLGGSSITWEHSIQMYTTMCKECLRLAPVANSRLDKETTPPDGRSLP 399  
QY 409 CGTELLIPILAVHDDQAIWGVNVEFNFARFADGVPPAAK-HPVGFIPEGLGYRTICGN 467  
DB 400 AGITVFIMWALHNHPYFW-EDPOVFNPLFRSR--ENSEKHHPYAFIPFSGJRNICIGH 456  
QY 468 LALIOAKTLTAVMIORFTEHAPYQAP 496  
DB 457 FALIECKVAVALTLIR--FKLAPDSRPP 483

RESULT 12  
AAU83606  
ID AAU83606 standard; protein; 505 AA.









CC cartilage disorders (e.g. sports injuries or arthritis), or wounds. The  
CC PRO polypeptides are useful in drug screening, particularly as targets  
CC for therapeutic intervention in these diseases, and in the diagnostic  
CC determination of the presence of these diseases. The PRO polypeptides are  
CC also useful as molecular weight markers, or for chromosome  
CC identification. The PRO genes are useful as hybridisation probes, or for  
CC screening libraries of human cDNA, genomic DNA or mRNA. The PRO genes may  
CC also be used in gene therapy, particularly for replacing a defective  
CC gene. This is the amino acid sequence of a novel human secreted and  
CC transmembrane PRO polypeptide

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Comugen Ltd.

OM protein - protein search, using sw model

Run on: May 19, 2004, 10:22:05 ; Search time 21 Seconds  
(without alignments)  
2381.885 Million cell updates/sec

Title: US-09-992-901-2  
Perfect score: 2730  
Sequence: 1 MEESSSMFIPKXVLVLSVIL.....LTPQHGAPITRRLLTNEED 520

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :  
1: PIR78:\*  
2: PIR2:\*  
3: PIR3:\*  
4: PIR4:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query %	Length	DB	ID	Description
1	2722	99.7	520	2	H84663	probable cytochrome
2	1113	40.8	524	2	T09999	cytochrome P450 -
3	1110	40.7	516	2	T10000	cytochrome P450 (C
4	1101	40.3	524	2	T09944	probable cytochrom
5	1028.5	37.7	505	2	D96781	cytochrome P450, p
6	1023	37.5	508	2	T16980	probable cytochrom
7	1019	37.3	517	2	T02192	cytochrome P450 ho
8	985.5	36.1	526	2	T02191	cytochrome P450 ho
9	943	34.5	518	2	T05876	Similar to Cytochr
10	900	33.0	476	2	D86306	hypothetical prote
11	887	32.5	512	2	A96695	laurate omega-hydr
12	523	19.2	509	1	O4RTLO	cytochrome P450 Cy
13	507	18.6	509	2	S47553	hypothetical prote
14	503	18.4	518	2	T24783	hypothetical prote
15	495	18.1	518	2	T24779	hypothetical prote
16	493	18.1	511	1	A34260	laurate omega-hydr
17	488.5	17.9	510	1	A29368	prostaglandin omeg
18	488.5	17.9	518	2	T20908	hypothetical prote
19	487.5	17.9	504	2	A60564	cytochrome P450 3A
20	486.5	17.8	517	2	T09067	hypothetical prote
21	484.5	17.7	519	2	I53015	fatty acid omega-h
22	482.5	17.7	507	2	JC7883	cytochrome P450 4X
23	482	17.7	537	2	JC4534	cytochrome P450 4F
24	481.5	17.6	519	2	UC0331	laurate omega-hydr
25	481	17.6	511	1	B40164	cytochrome P450 4B
26	480	17.6	519	2	T24784	hypothetical prote
27	479.5	17.6	504	2	S50892	cytochrome P450 3A
28	478	17.5	520	2	T24778	hypothetical prote
29	476	17.4	522	2	JC4532	cytochrome P450 4F

30	475	17.4	511	1	B34160	cytochrome P450 4A
31	475	17.4	544	2	T27750	hypothetical prote
32	474.5	17.4	504	2	A32965	cytochrome P450 4A
33	473	17.3	511	1	O4HQB1	cytochrome P450 4B
34	473	17.2	524	2	S29723	cytochrome P450 4F
35	470	17.2	534	2	T23376	hypothetical prote
36	467.5	17.1	520	2	T24777	hypothetical prote
37	466	17.1	515	2	T19576	hypothetical prote
38	466	17.1	524	2	UC7594	cytochrome P450 en
39	466	17.1	524	2	UC7598	cytochrome P450 en
40	465.5	17.1	520	2	T24780	hypothetical prote
41	464	17.0	511	2	S66472	cytochrome P450 4B
42	463.5	17.0	508	1	A36304	cytochrome P450 4A
43	463.5	17.0	507	1	A34160	laurate omega-hydr
44	463	17.0	507	1	A32966	cytochrome P450 4A
45	463	17.0	510	2	T19577	hypothetical prote

ALIGNMENTS

RESULT 1

H84663  
probable cytochrome P450 [imported] - Arabidopsis thaliana  
C.Species: Arabidopsis thaliana (mouse-ear cress)  
C.Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 02-Feb-2001  
C.Accession: H84663  
R.Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.  
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; Yencken, S.E.; Umayam, L.; Tallon,  
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter  
Nature 402, 761-768, 1999  
A.Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.  
A.Reference number: A84420; MUID:20083487; PMID:10617197  
A.Accession: H84663  
A.Status: preliminary  
A.Molecule type: DNA  
A.Residues: 1-520 <STO>  
A.Cross-References: GB:AE002093; NID:92760837; PIDN:AA95305.1; GSPDB:GN00139  
C.Genetics:  
A.Gene: At2g26710  
A.Map position: 2

Query Match	Score	DB 2	Length	520
Best local similarity	99.8%	Pred. No. 2.3e-197		
Matches	519	Conservative	0	Mismatches 1; Indels 0; Gaps 0;
QY	1	MEESSSMFIPKXVLVLSVILSVIVKMSLLMRPKIEHFKQIRGPPYHFGNYK	60	
DB	1	MEESSSMFIPKXVLVLSVILSVIVKMSLLMRPKIEHFKQIRGPPYHFGNYK	60	
QY	61	ELVGMMLKASHPMPESHNLPRVLSFYHMKKIYGATPLVWGGPFRLLTVADPDILREI	120	
DB	61	ELVGMMLKASHPMPESHNLPRVLSFYHMKKIYGATPLVWGGPFRLLTVADPDILREI	120	
QY	121	FSKSEPFYKQNAHVLQLEGDGLSLKSGKMAHHRKILSPFHMENLKLIVPVVLSVT	180	
DB	121	FSKSEPFYKQNAHVLQLEGDGLSLKSGKMAHHRKILSPFHMENLKLIVPVVLSVT	180	
QY	181	DMVDMKSDKLSSENGEVEVDVYEMFOILTEDEVISRTAFSGSYEDGRAVFRLOAQOMLLCAE	240	
DB	181	DMVDMKSDKLSSENGEVEVDVYEMFOILTEDEVISRTAFSGSYEDGRAVFRLOAQOMLLCAE	240	
QY	241	AFOCVFIPGVFFPTFRNLKSRKLDKEIRKSLKLIERRQNAIDEGECKEPPAAKDLL	300	
DB	241	AFOCVFIPGVFFPTFRNLKSRKLDKEIRKSLKLIERRQNAIDEGECKEPPAAKDLL	300	
QY	301	GLMTQAKNVTVQDIVECKSFPPAGKQTSNLTWTTLISMPBNQAKRDEVAVCGS	360	
DB	301	GLMTQAKNVTVQDIVECKSFPPAGKQTSNLTWTTLISMPBNQAKRDEVAVCGS	360	
QY	361	RDVPRKHVYKTLKLSMILNESLVPYVATIRBASDVKLGSGYKIPGCTELLIPILAV	420	
DB	361	RDVPRKHVYKTLKLSMILNESLVPYVATIRBASDVKLGSGYKIPGCTELLIPILAV	420	

QY 421 HHDOAIWGNVNEENPAPFADGVPRAAKHVGFIPFGLGVRTICGNLAILQAKTLAVM 480  
 DB 421 HHDOAIWGNVNEENPAPFADGVPRAAKHVGFIPFGLGVRTICGNLAILQAKTLAVM 480  
 QY 481 IORFTFHAPTYQHAPVLMILLYPHQGAFTFRRLTNHED 520  
 DB 481 IORFTFHAPTYQHAPVLMILLYPHQGAFTFRRLTNHED 520

## RESULT 2

T09999  
 cytochrome P450 - Madagascarc periwinkle  
 C/Species: Catharantus roseus (Madagascarc periwinkle)  
 C/Date: 16-Jul-1999 #sequence\_revision 16-Jul-1999 #text\_change 16-Feb-2001  
 C/Accession: T09999  
 R/Mangold, U.; Eichel, J.; Batschauer, A.; Lanz, T.; Kaiser, T.; Spangenberg, G.; Werck-  
 plant Sci. 96, 129-136, 1994  
 A/Title: Gene and cDNA for plant cytochrome P450 proteins (CYP72 family) from Catharantus  
 A/Reference number: Z16915  
 A/Accession: T09999  
 A/Status: preliminary; translated from GB/EMBL/DBJ  
 A/Molecule type: DNA  
 A/Residues: 1-524 <MAN>  
 A/Cross-references: EMBL:L19074; NID:G404687; PID:G404688  
 A/Experimental source: cv. cp3  
 C/Genetics: CYP72B  
 A/Intons: 96/1; 170/3; 252/2; 381/3  
 C/Superfamily: human cytochrome P450 CYP4B1; cytochrome P450 homology  
 C/Keywords: heme; iron; metalloprotein  
 F/329-492/Domain: cytochrome P450 homology <P45>  
 F/470/Binding site: heme iron (Cys) (axial ligand) #status predicted

Query Match 40.8%; Score 1113; DB 2; Length 524;  
 Best Local Similarity 41.6%; Pred. No. 4.7e-76;  
 Matches 216; Conservative 113; Mismatches 164; Indels 26; Gaps 8;

QY 15 VLSVLSIVYKMSL---WMPRKIEEHFSKQIGRPYHFFIGNVKELVGMKAS 71  
 DB 12 IATITFALVAMAMRVLDMAWFTPKRIKLRQGFNGNRYRLVGDVSGHMQEALS 71  
 QY 72 HMPFSNHLIPRVLSFYHMKRTIYGATFLVFGPTFRITVADPDILREISKEFEYKN- 130  
 DB 72 KMEFNDIVPLRMFINHTINTYGNSTWGRIPRLHWEPELKEVLTSSKYQNF 131  
 QY 131 EAH-PLVQLGDDGLSLKGEKAAHRRKISPTFMENLKLVPVLYKSTVDWYDKMSD 189  
 DB 132 DVANPLVLFLL-TGVSGFEGAKSKRRILISPAFTLEKLSMDFAPALCYHMDLTWEKI 190  
 QY 190 LSENGEVVDVYEWFOILTEDYISRTAFGSSYEDGRAVFLQAOQMLCAEAFQKPIPG 249  
 DB 191 AEKQSHVIDFPTFDVLTSDVYSKVAFGSTYEBGKIFRLKEIMDLTIDCKEDVYIPG 250  
 QY 250 YRFPFRGNLKSRLDKRIKSLKLEERRQNALDGESECKEPAKDLGLMIQAK-- 307  
 DB 251 WSLPTKRNKMKMEIKEL-TDMLRFTINKRKAL-----KAGEGDDLLGVLLBSNIQ 304  
 QY 308 -----NTVODIVECKSPFFAGKQTSNLLTWTTILLSMPEWQAKADEYLK 356  
 DB 305 EIQKQNRKDGMSINDVIEBGLFYFAGQETTVLLTWTTILLSKHEWQERAREVLTQ 364  
 QY 357 VCGSRDVTQKHVYKLTSLMILNESLRIPYIATIRAKSDVKGKYKIPCGTELLIP 416  
 DB 365 AFG-KMKPEFERLNLKTVSMILYEVRLYPVIDLTIKIHEDTKLGPYITIPACTQWMLP 423  
 QY 417 IIAVHDOAIWGNVNEENPAPFADGVPRAAKHVGFIPFGLGVRTICGNLAILQAKT 476  
 DB 424 TWMLREKSIWGEDAMENPAPFADGVANATKNNVTYLPFGWGRVLCIGNFALLQAKLG 483  
 QY 477 LAVMIQRTFHAPTYQHAPVLMILLYPHQGAFTFRRL 515  
 DB 484 LAMILQRTFDVAPSYVAHAPFTILTVOQFGSHVYKKL 522

## RESULT 3

T10000  
 cytochrome P450 (CYP72C) - Madagascarc periwinkle (fragment)  
 C/Species: Catharantus roseus (Madagascarc periwinkle)  
 C/Date: 16-Jul-1999 #sequence\_revision 16-Jul-1999 #text\_change 16-Feb-2001  
 C/Accession: T10000  
 R/Mangold, U.; Eichel, J.; Batschauer, A.; Lanz, T.; Kaiser, T.; Spangenberg, G.; Werck-  
 plant Sci. 96, 129-136, 1994  
 A/Title: Gene and cDNA for plant cytochrome P450 proteins (CYP72 family) from Catharar  
 A/Reference number: Z16915  
 A/Accession: T10000  
 A/Status: preliminary; translated from GB/EMBL/DBJ  
 A/Molecule type: mRNA  
 A/Residues: 1-516 <MAN>  
 A/Cross-references: EMBL:L19075; NID:G404689; PID:G404690  
 A/Experimental source: cv. cp3  
 C/Genetics: CYP72C  
 A/Status: unassigned cytochrome P450; cytochrome P450 homology  
 C/Superfamily: human cytochrome P450 CYP4B1; cytochrome P450 homology  
 C/Keywords: heme; iron; metalloprotein  
 F/318-481/Domain: cytochrome P450 homology <P45>  
 F/459/Binding site: heme iron (Cys) (axial ligand) #status predicted

Query Match 40.7%; Score 1110; DB 2; Length 516;  
 Best Local Similarity 41.8%; Pred. No. 7.8e-76;  
 Matches 217; Conservative 112; Mismatches 164; Indels 26; Gaps 8;

QY 15 VLSVLSIVYKMSL---WMPRKIEEHFSKQIGRPYHFFIGNVKELVGMKAS 71  
 DB 1 IATITFALVAMAMRVLDMAWFTPKRIKLRQGFNGNRYRLVGDVSGHMQEALS 60  
 QY 72 HMPFSNHLIPRVLSFYHMKRTIYGATFLVFGPTFRITVADPDILREISKEFEYKN- 130  
 DB 61 NPMEDNDIVPLRMFINHTINTYGNSTWGRIPRLHWEPELKEVLTSSKYQNF 120  
 QY 131 EAH-PLVQLGDDGLSLKGEKAAHRRKISPTFMENLKLVPVLYKSTVDWYDKMSD 189  
 DB 121 DVANPLVLFLL-TGVSGFEGAKSKRRILISPAFTLEKLSMDFAPALCYHMDLTWEKI 179  
 QY 190 LSENGEVVDVYEWFOILTEDYISRTAFGSSYEDGRAVFLQAOQMLCAEAFQKPIPG 249  
 DB 180 AEKQSHVIDFPTFDVLTSDVYSKVAFGSTYEBGKIFRLKEIMDLTIDCKEDVYIPG 239  
 QY 250 YRFPFRGNLKSRLDKRIKSLKLEERRQNALDGESECKEPAKDLGLMIQAK-- 307  
 DB 240 WSLPTKRNKMKMEIKEL-TDMLRFTINKRKAL-----KAGEGDDLLGVLLBSNIQ 293  
 QY 308 -----NTVODIVECKSPFFAGKQTSNLLTWTTILLSMPEWQAKADEYLK 356  
 DB 294 EIQKQNRKDGMSINDVIEBGLFYFAGQETTVLLTWTTILLSKHEWQERAREVLTQ 353  
 QY 357 VCGSRDVTQKHVYKLTSLMILNESLRIPYIATIRAKSDVKGKYKIPCGTELLIP 416  
 DB 354 AFG-KMKPEFERLNLKTVSMILYEVRLYPVIDLTIKIHEDTKLGPYITIPACTQWMLP 412  
 QY 417 IIAVHDOAIWGNVNEENPAPFADGVPRAAKHVGFIPFGLGVRTICGNLAILQAKT 476  
 DB 413 TWMLREKSIWGEDAMENPAPFADGVANATKNNVTYLPFGWGRVLCIGNFALLQAKLG 472  
 QY 477 LAVMIQRTFHAPTYQHAPVLMILLYPHQGAFTFRRL 515  
 DB 473 LAMILQRTFDVAPSYVAHAPFTILTVOQFGSHVYKKL 511

## RESULT 4

T09944  
 Probable cytochrome P450 protein - Madagascarc periwinkle  
 N/Alternate names: CYP72 protein  
 C/Species: Catharantus roseus (Madagascarc periwinkle)  
 C/Date: 16-Jul-1999 #sequence\_revision 16-Jul-1999 #text\_change 16-Feb-2001  
 C/Accession: T09944



```

Query Match      37.5%; Score 1023; DB 2; Length 508;
Best Local Similarity 40.2%; Pred. No. 2.8e-69;
Matches 205; Conservative 104; Mismatches 179; Indels 22; Gaps 9;

QY 16 LSVILSVIVYKGMILLWMRPKRIEHHFSKOGIRGPPHFIQNVKELVGMMLKASSHPMF 75
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 11 VGIIVAFPCGFELTITMWMPPKMIKKLKKSGIYVPRYKLLFGMLKEIMKSKXKQPL 69
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 76 FSNHILDPVLSFYHHMKYIGATPLVMFGPTFLTVADPDLIRLBITSKSFYKKNAPHL 135
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 70 LTHDIIIPVNPFLQHVANTYKKIFVLWMGPFPRTVTDPKLIRBLINRYTEHFKPEANAF 129
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 136 VKQJEGDGLSLKKEKXAHNRKIIISPFHEMLTLVPLVLSKVTMDVMSDKLSNGE 195
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 130 I-HFVLTGLASYDEKMDTRHKLINPAHYEKLKRMPPARAMCDEMLNNESSVGKTS 188
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 196 VEVDYVWFQILTIEDVISRTAFGSSYEDBGRVFLQAQOM-LICAEAFQVFTIPGRFFP 254
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 189 CEIDVLANFNLGSGVIRAFSGNIAEGRIIFLQKQCELLIASPF-TLFPFSLRFLP 247
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 255 TRGNLSRKLDKEIRKSLILKLIERRQNALIDGSEKEPAADULLGIMTQAKN----- 308
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 248 TASNRKAKYIEKKVSLIRGIIIEKREDAVRGISDN-----DILQILLKARADNKR 301
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 309 ---VIVQDIVEBECKSFPPAGKQTTSNLLTWTTLISMHPEWQAARDEVLRVCGSRDPT 365
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 302 AGGLTTEVEIECKEFYFAGQDTTTALISMTVVVLSHPEMDAKRADEVQVIG-KKKPK 360
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 366 KDHVVKLKLISMLINESRLYRPPIVATIRAKSDVKLGGYKIPCGTELLIPIAVHDOA 425
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 361 FDDLNLQKLMMNIPBEVLRILYPAIF-LIRNSKSKYKGDWMTIPAGVOVCVPTHLVHRDPE 419
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 426 IWMDVNEFNPNARPADGVPRAAKHPVGTIPRGLGVRITIGQNALILQAKLTIAIMIQRT 485
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 420 VMGDALLFNPRFSEGVSKAAEQM-YFPFGMGPRMCIQNFQMLEKULLISLIQRFW 478
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 486 FHIAPTYQHAPFTVMLLYPQAGAPITFRRL 515
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 479 FELSPSYTHAPLLTIMRPQAGQITVAKL 508
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

RESULT 7  
 T02192  
 probable cytochrome P450 At2g46950 [imported] - Arabidopsis thaliana  
 N|Alternate names: cytochrome P450 homolog F14M4.22  
 C|Species: Arabidopsis thaliana (mouse-ear cress)  
 C|Date: 05-Mar-1999 #sequence\_revision 05-Mar-1999 #text\_change 16-Feb-2001  
 C|Accession: T02192 C84909  
 R|Rounsley, S.D.; Lin, X.; Kaul, S.; Shea, T.P.; Fujii, C.Y.; Mason, T.M.; Shen, M.; Rott  
 Submitted to the EMBL Data Library, September 1998  
 A|Description: Arabidopsis thaliana chromosome II BAC F14M4 genomic sequence.  
 A|Reference number: Z14609  
 A|Accession: T02192  
 A|Status: translated from GB/EMBL/DBJ  
 A|Molecule type: DNA  
 A|Residues: 1-517 <R0U>  
 A|Cross-references: EMBL:AC004411; NID:g3522932; PID:g3522846  
 A|Experimental source: cultivar Columbia  
 R|Ritt, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Bentto, M.I.; Town, C.D.; Fujii, C.Y.;  
 M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.  
 uens, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Frazer, C.M.; Venter, A.  
 Nature 402, 761-768, 1999  
 A|Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.  
 A|Reference number: A84420; M|ID:20083487; F|MD:1061197  
 A|Accession: C84909  
 A|Status: preliminary  
 A|Molecule type: DNA  
 A|Residues: 1-517 <R0U>  
 A|Cross-references: GB:AE002093; NID:g3522946; PIDN:AAC34228.1; GSFDB:GN00139  
 C|Genetics:  
 A|Gene: At2g46950, F14M4.22  
 A|Map position: 2  
 A|Introns: 95/1; 169/3; 252/2; 374/3

C/Superfamily: human cytochrome P450 CYPB1; cytochrome P450 homology	
C/Keywords: heme; iron; metalloprotein	
F.321-485/Domain: cytochrome P450 homology <P45-	
F.463/Binding site: heme iron (Cys) (axial ligand) #status predicted	
Query Match:	37.3%; Score 10.9; DB 2; Length 517;
Best Local Similarity	35.8%; Pred. No. 5.7e-68;
Matches 204; Conservative 108; Mismatches 176; Indels 24; Gaps 8;	

[illegible]

RESULT 8  
 T02191  
 cytochrome P450 homolog F14M4.21 - Arabidopsis thaliana  
 C:/Species: Arabidopsis thaliana (mouse-ear cress)  
 C:/Date: 05-Mar-1999 #sequence\_revision 05-Mar-1999 #text\_change 16-Feb-2001  
 C:/Accession: T02191, D84909  
 R:/Rummenley, S.D.; Lin, X.; Kaul, S.; Shea, T.P.; Fujii, C.Y.; Mason, T.M.; Sher, M.;  
 submitted to the EMBL Data Library, September 1998  
 A:/Description: Arabidopsis thaliana chromosome II BAC F14M4 genomic sequence.  
 A:/Reference number: Z14609  
 A:/Accession: T02191  
 A:/Status: translated from GB/EMBL/DBJ  
 A:/Molecule type: DNA  
 A:/Residues: 1-526 <ROU>  
 A:/Cross-references: EMBL:AC004411; NID:G3522932; PID:G3522945  
 A:/Experimental source: cultivar Columbia  
 R:/Lin, X.; Kaul, S.; Rummenley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;  
 M.; Koo, H.; Mofatt, K.S.; Cronin, L.A.; Shen, M.; Vankken, S.E.; Umayam, L.; Tallon  
 eus, D.; Nleman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Ventre  
 Nature 402, 761-768, 1999  
 A:/Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.  
 A:/Reference number: A84420; WUID:20083487; PMID:1061197  
 A:/Accession: D84909  
 A:/Status: preliminary  
 A:/Molecule type: DNA  
 A:/Residues: 1-526 <ROU>  
 A:/Cross-references: GB:AE002093; NID:G3522945; PID:N.AAC34227.1; GSPDB:IGN00139

C:Genetics:  
 A:Gene: At2g46960, F14M4.21  
 A:Map position: 2  
 A:Insertions: 93/1; 168/3; 254/2; 375/3  
 C:Superfamily: human cytochrome P450 CYP4B1; cytochrome P450 homology  
 C:Keywords: heme; iron; metalloprotein  
 F:322-493/Domain: cytochrome P450 homology <P45>  
 F:471/Binding site: heme iron (Cys) (axial ligand) #status predicted

Query Match 36.1%; Score 985.5; DB 2; Length 526;  
 Best Local Similarity 37.2%; Pred. No. 2e-66;  
 Matches 196; Conservative 117; Mismatches 179; Indels 35; Gaps 9;

```

QY 13 VLVLSVLSLVYKGSILMKMRPKIEHPSKQIGRPFYHFGNVGLVGMKASSH 72
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
D 9 LVLVLILGIRKAFMILVWHPVLTIRRLKQGISGPNYRIFYGSLSEIKKM--KRESH 66
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 73 PM--PFSNHLPRVLSFYHMRKIYGATFLVWFGPTFLTVADPLIREIFS--KSEFE 128
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
D 67 LSLIDPSSNDIFRILPHYQKMSQYGETFLVWNGTERPICSDBLAKTMLSNKLGFV 126
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 129 KNEAHP-LVKQLSGDGLSLKGEKMAHKKIISPTFMENKLLVPPVVLKSVTMDVKM- 186
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
D 127 KSKARPEAVLVGSKGLVFEGADWVRHRIINPAFSIDRLKIMTVWDCTLKMLEBR 186
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 187 ---SDKLSNGEVEVDVYEWFOILTEDVTSRTAFGSYEDGRAVFLQAOQMLCAEAFQ 243
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
D 187 KESTKETEPKIKKEMNEFORLTADILATSAFGSYVEGLEVRSGMELKRCYTSLSN 246
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 244 KVPFPGYRFPPTGNLKSRLDKERKSLDKLIERRONAIDEGECKEPAKDLGLM 303
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
D 247 QVSLPGYQLPTPSNIRVWMLERKMDNSIKRILSRISLQ8-----KSDYGDLLGL 297
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 304 IQAKN-----VTQDIVECKSFPPAKQOTSNLITWTLLSMHPBWKAKDEYLR 356
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
D 298 LKAYNTEGKERKMSIEIITHECTFFFGHETTSNLTAVTMLSLSHQWCKLEELFK 357
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 357 VCGSRDVPYTKDHYVKK-----TSMILNESLRYPPVATIRAKSDVGLGYSKIPC 409
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
D 358 ECGEKETPDETSKTLKRPVFLQNMVIMESLRLYGVSLARASVNIKGLDELIFK 417
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 410 GTELLPIITIAVHDDQAIWGDVNEFNPAPADGVPRAAGPVGFIPLGIVATTCGNTLA 469
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
D 418 GTLVVIPLLMGHSKTLMGSDAKFPMRPANVSAAAHNALLAFSGVPACTIGQNFV 477
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 470 ILQAKTLAVMIQRFPR-HLAPYOHAPTVMLLYPQHGAPITFRRL 515
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
D 478 MIEAKTVLMILQRFRIISLCEYKHTPVDNVITIGQYGLPWLQPL 524
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

RESULT 9  
 T05876  
 cytochrome P450 homolog T29A15.200 - Arabidopsis thaliana  
 C:Species: Arabidopsis thaliana (mouse-ear cress)  
 C:Date: 09-Apr-1999 #sequence\_revision 09-Apr-1999 #text\_change 28-Jul-2000  
 C:Accession: T05876  
 R:Bevan, M.; Van Der Schueren, J.; Chuang, Y.-J.; Voet, M.; Robben, J.; Volckaert, G.; He submitted to the Protein Sequence Database, March 1999  
 A:Reference number: Z15455  
 A:Accession: T05876  
 A:Molecule type: DNA  
 A:Residues: 1-518 <BRV>  
 A:Cross-references: EMBL:AL035602; GSPDB:GN00062; ATSP:T29A15.200  
 C:Genetics:  
 A:Gene: ATSP:T29A15.200  
 A:Map position: 4  
 A:Insertions: 95/1; 169/3; 252/2; 376/3  
 C:Superfamily: human cytochrome P450 CYP4B1; cytochrome P450 homology  
 C:Keywords: heme; iron; metalloprotein  
 F:323-467/Domain: cytochrome P450 homology <P45>  
 F:465/Binding site: heme iron (Cys) (axial ligand) #status predicted

Query Match 34.5%; Score 943; DB 2; Length 518;  
 Best Local Similarity 37.5%; Pred. No. 3.1e-63;  
 Matches 196; Conservative 107; Mismatches 195; Indels 24; Gaps 8;

```

QY 10 IPKVLVLSVLSLVY---KMSLLMRPKIEHPSKQIGRPFYHFGNVGLVGM 65
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
D 4 ISTINLTIVLLFVSKIKWACWILLRPLMLSRFKQGISGPKYKILYGNISEIKM 63
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 66 MKASHPM-PFSNHLPRVLSFYHMRKIYGATFLVWFGPTFLTVADPLIREIFS-SK 123
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
D 64 KKEADLCVDPSPNDIFPRVPQYQMSQYGDITLFTGTGYPTIYNHBLAKVLSK 123
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 124 SEFEKNEAHPVYKQLSGDGLSLKGEKMAHKKIISPTFMENKLLVPPVVLKSVTMDV 183
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
D 124 FGTITIPVRRPFFILFGKLSFIQDDKIRRRRLINPAFSMDRLKATQPMGDCITLRF 183
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 184 DKMSDLSNGEV--EVDYEWFOILTEDVTSRTAFGSYEDGRAVFLQAOQMLCAEA 241
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
D 184 EEMR-KORNGEVLKIEISKEFHRLTDIITAFSGSYAEGLELCSQTELEKYISS 242
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 242 FQKVFIPGYRFPPTGNLKSRLDKERKSLDKLIERRONAIDEGECKEPA-AKDL 300
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
D 243 LTVVIFPGYQLPTPNLKLMLHKKVKNISRIIDSLKS-----KCKTYGDDLL 295
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 301 GLMTQA-----KNVTQDIVECKSFPPAKQOTSNLITWTLLSMHPBWKAKDE 353
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
D 296 GVMLTAKSNEYRKKRMDIEECKNFYAGQYTSLLITWTMLSLSHQWCKLEELFK 355
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 354 VARCGSRDVPYTKDHYVKKITSMILNESLRYPPVATIRAKSDVGLGYSKIPC 413
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
D 356 VENECGRDXIPDTDFSKLNMVIMESLRLYGVVITISREATQMKVGHLEIPKYSI 415
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 414 LIPIIAVHDDQAIWGDVNEFNPAPADGVPRAAGPVGFIPLGIVATTCGNTLAIIQA 473
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
D 416 IIPLLMGHDKAIWGDADQFPLRPFENGISQATIHFNALLPFSIGPACTIGQNFV 475
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 474 KTLAVMIQRFPHLAPYOHAPTVMLLYPQHGAPITFRRL 515
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
D 476 KTVLTMILQFQLSPEYKHTPVDHFLFPQYGLPWLMLPL 517
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

RESULT 10  
 D86306  
 Similar to Cytochrome P450 [imported] - Arabidopsis thaliana  
 C:Species: Arabidopsis thaliana (mouse-ear cress)  
 C:Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 18-Nov-2002  
 C:Accession: D86306  
 R:Theologian, A.; Ecker, J.R.; Palm, C.D.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso Chn, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K ansen, N.F.; Hughes, B.; Hutzler, L.  
 Nature 408, 816-820, 2000  
 A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luos, J.S.; Maiti, R.; Marziani Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.  
 A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tabor ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.  
 A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.  
 A:Reference number: A86141; MIDB:21016719; PMID:11130712  
 A:Accession: D86306  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-476 <STO>  
 A:Cross-references: GB:AE005172; NID:95734759; PIDN:AA05024.1; GSPDB:GN00141  
 C:Genetics:  
 A:Map position: 1  
 C:Superfamily: Bacillus halodurans cytochrome P450 BH0579; cytochrome P450 homology  
 C:Keywords: heme; iron; metalloprotein  
 F:424/Binding site: heme iron (Cys) (axial ligand) #status predicted

Query Match 33.0%; Score 900; DB 2; Length 476;  
 Best Local Similarity 35.2%; Pred. No. 4.7e-60;  
 Matches 180; Conservative 107; Mismatches 177; Indels 48; Gaps 5;



QY 10 IPKVLVLSVILSV--IVKMSLLMWRPKIEHFSKQIRGPPYHFIQNVKLVGMML 67  
 DB 7 VRKVLFIQFLLILNMVRAVNMVWLRKFLKLEKTKQSGNSYRLIMQDMKESNQMO 66  
 QY 68 KASSHPMPFSHNILPRVLSFYHNRKITYGATFLVPGPTFLVADBDLIRELFSSEY 127  
 DB 67 VAHSLPLPLDADFLPRMPLFHHVVLKQKCKFTWYPRVNVMPETLREIMSGHELF 126  
 QY 128 EK----NEAHPLVKQLEDGDLISLKGEMAHHRKIISPTFPMENLKLVVVLKSYTDVY 183  
 DB 127 PPKPKGSHNHVFL-----SGLLNHEGPKMSKHSILNPARIDLKSILFAFNSCKEML 181  
 QY 184 DKMSDKLSENGEVDVYEMFOILTDEVISTRAGSSYEDGRAVFLQAOQMLCAEAQ 243  
 DB 182 EEMERLASAKGTMEJDSWTHCHDLTRNLARASFGSDYKQGIKIFELIQEQIDLGILAR 241  
 QY 244 KYEIPGYRFPPTRGNLKSKLDKELRKSLLKLIERRONAIEDGECEKPEAKDILGLM 303  
 DB 242 AVTIGSKFLPTKFRRLRRETRBDRKAMFKMITKESEIKRGR----- 285  
 QY 304 IQAKNVTVODIVECKSFPPAGKQTSNLTWTITLLSMPEWQAKARDEVLCVGSRDV 363  
 DB 286 -----AGQVTSLSLFWTLVALSOHQDMQNKARDEISOAFGNR- 324  
 QY 364 PPKDHYVVKLTLSMLINESLRLYPVATIRAKSDYKLGSKYKIPGTELLIPLVHND 423  
 DB 325 PPEGLSHLKVATMLLHEVRLYSPAYTCITQOEYKLEBFSLPEGVVVTIPLLVHND 384  
 QY 424 QAINQDVNEFNPAFADGVPPAAKHPVGFIPFGIGVRTCIGQNLAILQAKTLAVMIO 483  
 DB 385 SOLWDDVYKPEKPEFANGVAGATKGRSLFLPFSSGPRTCIGQFSLQAKTLAVMIO 444  
 QY 484 FTFHLAPTYQAHPTVLMMLYPQHGAPITFRRL 515  
 DB 445 FSEVLSPSYTHAPFPAAFTTFQHGAMLIIRKL 476

## RESULT 11

A96695  
 hypothetical protein P5A8.3 [imported] - Arabidopsis thaliana  
 C/Species: Arabidopsis thaliana (mouse-ear cress)  
 C/Date: 02-Mar-2001 #sequence\_rev1stion 02-Mar-2001 #text\_change 20-Apr-2001  
 C/accession: A96695  
 R/theologian: A.; Ecker, J.R.; Palm, C.J.; Federpiet, N.A.; Kaul, S.; White, O.; Alonso,  
 Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;  
 anen, N.F.; Hughes, B.; Huizar, L.  
 Nature 408, 816-820, 2000  
 A/authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.  
 C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Malt, R.; Marziani,  
 Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.  
 A/authors: Salzberg, S.L.; Schwartz, J.R.; Shim, P.; Southwick, A.M.; Sun, H.; Tallon,  
 ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.  
 A/title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.  
 A/reference number: A86141; MUID:21016719; PMID:11307072  
 A/accession: A96695  
 A/status: preliminary  
 A/molecule type: DNA  
 A/residues: 1-512 <STO>  
 C/cross-references: GB:A005173; NID:g4204278; PIDN:AA010659.1; GSPDB:GN00141  
 A/cene: P5A8.3  
 A/map position: 1  
 A/superfamily: human cytochrome P450 CYP4B1, cytochrome P450 homology  
 C/keywords: heme; iron; metalloprotein  
 F/458/Binding site: heme iron (axial ligand) #status predicted

Query Match 32.5%; Score 887; DB 2; Length 512;  
 Best Local Similarity 35.6%; Pred No. 5e-59;  
 Matches 185; Conservative 109; Mismatches 189; Indels 36; Gaps 8;  
 QY 13 VLVLSVLSVIVKMSLLMWRPKIEHFSKQIRGPPYHFIQNVKLVGMML-KASS 71  
 DB 12 VLVMTILR-VLYDSICCYFLPRRIKEMERGIGTGPRLLTGNIIIDISKMLSHSASN 70

QY 72 HPMPSHNILPRVLSFYHNRKITYGATFLVPGPTFLVADBDLIRELFSKSEYEKNE 131  
 DB 71 DCSISHNIVPRLLPHVVSMSKQYGRFIMNGNTEPRLCLTEEMIKELLTGH----- 123  
 QY 132 AHPL-----VKQLEDGDLISLKGEMAHHRKIISPTFPMENLKLVVVLKSYTD 181  
 DB 124 -NPVTGSKWLQOQSTKFFIRGLLMANGEAMHQRMAAPAFRDLKGYAAMVECTM 182  
 QY 182 MVDKMSDKLSENGEVDVYEMFOILTDEVISTRAGSSYEDGRAVFLQAOQMLCAEA 241  
 DB 183 MAERLKEVGE-----EVEIEENRRLADLISNTEGSSGDKKELFSLTLYLQRLCAQ 238  
 QY 242 FQKFTIYGRFPPTRGNLKSKLDKELRKSLLKLIERRONAIEDGECEKPEAKDILG 301  
 DB 239 TRHLCFPGSRFLPSKVRREIKSLTEVERLMBEIDSRKDSVEIGRSSSYGD---DLIG 294  
 QY 302 LMI-----QAKNVTVODIVECKSFPPAGKQTSNLTWTITLLSMPEWQAKARDEVLR 356  
 DB 295 LILNQMSNKNANVQVIMDECKTFFGTGHTSLTLTLMLANFTMQDNVDRVQ 354  
 QY 357 VCGSRDVPYKDHVVKLTLSMLINESLRLYPVATIRAKSDYKLGSKYKIPGTELLP 416  
 DB 355 VCGQDGVPSVEQLSLSLTKVINESLRLYPATLDPRAFDIKLGDLIIPKGLSIWIP 414  
 QY 417 IIAVHDOAIVQNDVNEFNPAFADGVPPAAKHPVGFIPFGIGVRTCIGQNLAILQAKLT 476  
 DB 415 VLAIHNSNELMGEANEFNBERFTTSFASR----FMPPAAGPRCIGQTFAMMEAKII 471  
 QY 477 LAVMIQRTFHLAPTYQAHPTVLMMLYPQHGAPITFRRL 515  
 DB 472 LAMVSKFSFAISENYAHAPVILTIKPYGVQLVKPL 510

## RESULT 12

OAKTIO  
 laurate omega-hydroxylase (EC 1.14.15.3) cytochrome P450 4A1 - rat  
 N/Alternate names: cytochrome P450 1A-omega; cytochrome P452  
 C/Species: Rattus norvegicus (Norway rat)  
 C/Date: 30-Sep-1991 #sequence\_rev1stion 30-Sep-1991 #text\_change 03-Mar-2000  
 C/accession: S01336; A26137; B32965; S21711  
 R/Barnshaw, D.; Dale, J.W.; Goldfarb, P.S.; Gibson, G.G.  
 FEBS Lett. 236, 357-361, 1988  
 A/title: Differential splicing in the 3'-non-coding region of rat cytochrome P-452 (P4  
 A/reference number: S01336; MUID:88312998; PMID:3410047  
 A/accession: S01336  
 A/molecule type: mRNA  
 A/residues: 1-509 <ENR>  
 A/cross-references: EMBL:X07259; NID:g56046; PIDN:CAA30245.1; PID:g56047  
 R/Harwick, J.F.; Song, B.U.; Huberman, E.; Gonzalez, F.J.  
 J. Biol. Chem. 262, 801-810, 1987  
 A/title: Isolation, complementary DNA sequence, and regulation of rat hepatic lauric  
 A/reference number: A26137; MUID:87109183; PMID:3027069  
 A/accession: A26137  
 A/molecule type: mRNA  
 A/residues: 1-509 <HAR>  
 A/cross-references: GB:M4972; NID:g203865; PIDN:AAA41061.1; PID:g203866  
 R/Kimura, S.; Hanloka, N.; Matsunaga, E.; Gonzalez, F.J.  
 DNA 9, 503-516, 1989  
 A/title: The rat clofibrate-inducible CYP4A gene subfamily I. Complete intron and exo  
 am element.  
 A/reference number: A32965; MUID:89356271; PMID:2766932  
 A/accession: B32965  
 A/molecule type: DNA  
 A/residues: 1-340, E, 342-509 <KIM>  
 A/cross-references: EMBL:M57718; NID:g203786; PIDN:AAA41038.1; PID:g203787  
 R/Okita, R.T.; Okita, J.R.  
 Arch. Biochem. Biophys. 294, 475-481, 1992  
 A/title: Characterization of a cytochrome P450 from Di-(2-ethylhexyl) phthalate-treat  
 A/reference number: S21711; MUID:92231570; PMID:1567203  
 A/accession: S21711  
 A/status: preliminary  
 A/molecule type: protein

A:Residues: 1-22 <OK>  
 A:Experimental source: strain Sprague-Dawley  
 A>Note: this sequence was confirmed by protein sequencing  
 C:Genetics:  
 A:Gene: CYP4A1  
 A:introns: 63/3; 112/1; 127/1; 169/3; 211/2; 263/1; 298/3; 362/2; 407/1; 428/3; 454/2  
 C:Superfamily: human cytochrome P450 CYP4B1; cytochrome P450 homology  
 C:Keywords: chromoprotein; electron transfer; endoplasmic reticulum; fatty acid oxidatio  
 F:317-478/Domain: cytochrome P450 homology <P45>  
 F:456/Binding site: heme iron (Cys) (axial ligand) #status predicted

Query Match 19.2%; Score 523; DB 1; Length 509;  
 Best Local Similarity 28.4%; Pred. No. 1.4e-31;  
 Matches 151; Conservative 105; Mismatches 218; Indels 58; Gaps 17;

```

QY 5 SSSWFIPKVLVLSVLIVYKMSLLMWRPKLIEHFSKQIGRPYHFTGNV-----59
DB 14 SISGFLQVASVIGLL--LLYKAVQFYLRQWLK--AFQFPSPFPFHFQHKQFGD 68
QY 60 KELVGMMLKASHMPFESHNLPRVLSFYHMRKIYGATFLVWF-GPTFRLTVADPDILR 118
DB 69 KELQOIMTCVZNPFSAF-----PR-----WFGSKAVLYVDDPMK 105
QY 119 EIFSKSEFEKNEAHPVYKQLEGGDLSLKEKMAHRRKIIPTFMENIKLIIVPVVLS 178
DB 106 VILGRSD-PKANGVYRLAPWIGYGLLLNGQPFQHRRLTPAFHYDLIKPYVKNMADS 164
QY 179 VTDMDVKMSDKLSENGEVVDVYEMFQILTEDVISRTAF--GSSYEDGRAVRLQAQD 235
DB 165 IRLMDKWEQIAGDSSIE--TFQHSIMLTDVTKCAFSHGVSQVDGNYKSYQAIQD 222
QY 236 L--LCAEAFQKVFIPGYAF--FPTRGMLKSR--KLDEIKRSLKLIERRQNAIDGEG 289
DB 223 LNDLFHSRVNRNIFHQNDITYFSSNGHLFNRACQALADHTDGVYLRKDQLQA--GELE 280
QY 290 ECKEPAKDLGLMTQAK-----NTVQDIVCEKSFPPAGKQTSNLTWTTILSMRP 344
DB 281 KYKKRRRLDFLDILLFARMENGSDLSKDLAEVDITFMFGHDTTASGVSWITVYALATHP 340
QY 345 EMOAKARDEVLRVCGSRDPVTKDHYVVKLTLSMILNESLRTPVATIRRAKSDVKL-G 403
DB 341 KHQQRCEBVSVLGDSSITMDHDOIPTTMCIKALNLPVPGLVRELSTVTFPD 400
QY 404 GYKIPCGTELLIPITAVHDDQAIWGDVNEPNRPADGVPPRAKGVGTFPGGLGYTC 463
DB 401 GRSLPKGIQVTLSTIYGLHNPKWPNP-EVPDPSPRFAPDSPR--HSHSFLPFGSGARRC 456
QY 464 IGQNLAILOAKTLTAVMIOQFTFHLAPTYGHAFTVLMILYPOHGAPITFRL 515
DB 457 IGKQPMSEMKVYVLTLLRFELLPDPTVPRVPLVSKNGIYLYKKL 508

```

## RESULT 13

S47553

cytochrome P450 Cyp4a - mouse

C/Species: Mus musculus (house mouse)

C/Date: 07-May-1995 #sequence\_revision 21-Jul-1995 #text\_change 04-Mar-2000

C/Accession: S47553

R:Henderson, C.J.; Bammler, T.; Wolf, C.R.

Biochim. Biophys. Acta 1200, 182-199, 1994

A:Title: Dedicated amino acid sequence of a murine cytochrome P-450 Cyp4a protein: develop

A:Reference number: S47553; PMID:94304927; PMID:8031839

A:Accession: S47553

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-509 &lt;HEN&gt;

A:Cross-references: GB:K69296; NID:9312695

A&gt;Note: this ORF is not annotated in GenBank entry MGCYP4A, release 109.0

C:Superfamily: human cytochrome P450 CYP4B1; cytochrome P450 homology

C:Keywords: chromoprotein; heme; iron; metalloprotein; transmembrane protein

F:317-478/Domain: cytochrome P450 homology &lt;P45&gt;

F:456/Binding site: heme iron (Cys) (axial ligand) #status predicted

F:456/Binding site: heme iron (Cys) (axial ligand) #status predicted

Query Match 18.6%; Score 507; DB 2; Length 509;  
 Best Local Similarity 28.0%; Pred. No. 2.2e-30;  
 Matches 149; Conservative 107; Mismatches 219; Indels 58; Gaps 17;

```

QY 4 ESSSWFIPKVLVLSVLIVYKMSLLMWRPKLIEHFSKQIGRPYHFTGNV---60
DB 13 DSISGFLQVASVIGLL--LLYKAVQFYLRQWLK--AFQFPSPFPFHFQHKQFGD 67
QY 61 --ELVGMMLKASHMPFESHNLPRVLSFYHMRKIYGATFLVWF-GPTFRLTVADPDIL 117
DB 68 DQELQEVSCIEPPSAF-----PR-----WFGSKAVLYVDDPM 104
QY 118 REIFSKSEFEKNEAHPVYKQLEGGDLSLKEKMAHRRKIIPTFMENIKLIIVPVVIL 177
DB 105 KVLGRSD-PKANGVYRLAPWIGYGLLLNGQPFQHRRLTPAFHYDLIKPYVKNMAD 163
QY 178 VTDMDVKMSDKLSENGEVVDVYEMFQILTEDVISRTAF--GSSYEDGRAVRLQAQD 234
DB 164 SIRLMDKM-ERLADQ-DSSIEIFQHSIMLTDVTKCAFSHGVSQVDGNYKSYQAIQD 221
QY 235 ML--LCAEAFQKVFIPGYAF--FPTRGMLKSR--LDKIRKSLKLIERRQNAIDGEG 288
DB 222 DNLNLFHSRVNRNIFHQNDITYKLSNGRLAKQACQALADHTDGVYLRKDQLQA--RKDQLQDEGEL 279
QY 289 ECKEPAKDLGLMTQAK-----NTVQDIVCEKSFPPAGKQTSNLTWTTILSMW 343
DB 280 EKIKRRRLDFLDILLFARMENGSDLSKDLAEVDITFMFGHDTTASGVSWITVYALATH 339
QY 344 PEMQAKARDEVLRVCGSRDPVTKDHYVVKLTLSMILNESLRTPVATIRRAKSDVKL- 402
DB 340 PDHQQRCEBVSVLGDSSITMDHDOIPTTMCIKALNLPVPGLVRELSTVTFPD 399
QY 403 GYKIPCGTELLIPITAVHDDQAIWGDVNEPNRPADGVPPRAKGVGTFPGGLGYTC 462
DB 400 DGRSLPKGVQVTLSTIYGLHNPKWPNP-EVPDPSPRFAPDSPR--HSHSFLPFGSGARRN 455
QY 463 CIGQNLAILOAKTLTAVMIOQFTFHLAPTYGHAFTVLMILYPOHGAPITFRL 515
DB 456 CIGQPMSEMKVYVLTLLRFELLPDPTVPRVPLVSKNGIYLYKKL 508

```

## RESULT 14

T24783

hypothetical protein T10B9.10 - Caenorhabditis elegans

C/Species: Caenorhabditis elegans

C/Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 16-Feb-2001

C/Accession: T24783

R:Gardner, A.

submitted to the EMBL Data Library, March 1995

A:Reference number: Z19935

A:Accession: T24783

A:Status: preliminary; translated from GB/EMBL/DBD

A:Molecule type: DNA

A:Residues: 1-518 &lt;WIL&gt;

A:Cross-references: EMBL:Z48717; PIDN:CAA88609.1; GSPDB:GN00020; CESP:T10B9.10

A:Experimental source: clone T10B9

C:Genetics:

A:Gene: CESP:T10B9.10

A:Map position: 2

A:introns: 62/3; 173/2; 206/3; 392/2; 440/2

C:Superfamily: human cytochrome P450 CYP2A5; cytochrome P450 homology

C:Keywords: heme; iron; metalloprotein

F:323-486/Domain: cytochrome P450 homology &lt;P45&gt;

F:464/Binding site: heme iron (Cys) (axial ligand) #status predicted

Query Match 18.4%; Score 503; DB 2; Length 518;  
 Best Local Similarity 26.6%; Pred. No. 4.6e-30;  
 Matches 148; Conservative 109; Mismatches 180; Indels 120; Gaps 20;

```

QY 17 SVLSVLIVYKQ-MSLLMWRPKLIEHFSKQIGRPYHFTGNVYKELVGMMLKASHMP 75
DB 4 SILAIALFVGIIISYVL-----IWSFWIRKGVKGRGLPFLGVYHK-----45

```

QY 76 FSHNLLPVLSEFYHMKRYGATFLVWFSGPTFLVADPDLIRIFSK--SEFY----- 127  
 DB 46 FTNENGAALKF--SEWTKKIGPVYIGTEGVKTLVSDPEFVVEVVKQFDFYGRKTLA 104  
 QY 128 -----EKNEAPLVKOLEGGLSLKGEKMAHHRKIIISPTFMENIKLIVPVVLSYVDM 182  
 DB 105 IQGDPNKKRVPLV-----AAQGRWKRRLRTLASPTFSNKSIRKIVGESVTEL 155  
 QY 183 VDKMSDKLSENGEVVDVYEWFPQILFEDVYSRTAFSS-----YEDGRAVFLQAO 233  
 DB 156 V-RSLERKASAEKKT-LDMLEYEQEFTWDIIGKAMGEEKSLMRANPMDKVTIIEGEN 213  
 QY 224 QMLCAEAFQKVFIPGYRFFPTGNLK-SRLDKKIRKSLIKLIERRON--AIDGEE 289  
 DB 214 NVFMTSGIFPVGALRNIFAKFPRLQMATDIOGILEKALKRLEOREDEKAGIEPSE 273  
 QY 290 ECKEPAKDILGLMIQA-----KNVTQDIVECKSFEPAG 325  
 DB 274 -----PODFIDLFLDASTVDPEGEAEODPAKSEVLKVDKRLTDEIIGQLFVLLAG 327  
 QY 326 KOTSNLLTWTTILSNHPEWQAKARDEVLRVCGSRDVPFKDHVVKLKTLSMILNES 385  
 DB 328 YDTTALSLSYATVLAHPEIQKLOEVDRECPDEPV--TFDQSLKYLECVKKEALRL 386  
 QY 386 YPPIVAT--RRAKSDVKLGKIKIPGTELLIPTIAVHDOAIKGDVNEFPAR----- 438  
 DB 387 YP--LASLVNHRKCLKTNNVLMGEIENGTINVDWLSLHDPKVGODVNEFPKESG 444  
 QY 439 ----FADGVPRAAKHPVGFIPGIGVATCIGOMLAILQAKLTIAVNIQRTFLHAF---- 490  
 DB 445 DELFPAK-----CYLPFGMGPRICIGRLAMEMKMLTLNLIKVTFTTETVYI 495  
 QY 491 ----TYQHAFTVLM 501  
 DB 496 PLKIVGTATIASSVTL 512

RESULT 15  
 T24779  
 hypothetical protein T10B9.3 - Caenorhabditis elegans  
 C/Species: Caenorhabditis elegans  
 C/Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 16-Feb-2001  
 C/Accession: T24779  
 R/Gardner, A.  
 submitted to the EMBL Data Library, March 1995  
 A/Reference number: Z19935  
 A/Accession: T24779  
 A/Status: preliminary; translated from GB/EMBL/DBJ  
 A/Molecule type: DNA  
 A/Residues: 1-518 <MIL>  
 A/Cross-references: EMBL:Z48717; PIDN:CAA8605.1; GSPDB:GN00020; CESP:T10B9.3  
 A/Experimental source: clone T10B9  
 C/Genetics:  
 A/Gene: CESP:T10B9.3  
 A/Map position: 2  
 A/Intons: 60/3; 171/2; 204/3; 390/2; 438/2  
 C/Superfamily: human cytochrome P450 CYP3A5; cytochrome P450 homology  
 C/Keywords: heme; iron; metalloprotein  
 F/321-485/Domain: cytochrome P450 homology <P45>  
 F/463/Binding site: heme iron (axial ligand) #status predicted

Query Match 18.1%; Score 495; DB 2; Length 518;  
 Best Local Similarity 26.7%; Pred. No. 1.8e-29;  
 Matches 142; Conservative 101; Mismatches 197; Indels 92; Gaps 17;

QY 15 VLSVILSYIVKMSLIMRPRKIEHSEKQIRGPPYHFFIGNVKELVGMILKASSHPM 74  
 DB 1 MIFVLISAVALLGVFTSV-----LMSYIRKRIKIP--RGPPG-----IGMLIQTLIDHEN 49  
 QY 75 PFSNHLIPVLSPFYHMKRYGATFLVWFSGPTFLVADPDLIRIFSK--SEFYEKNEA 132  
 DB 50 P-----PFLKYRDMTKQYGVYGFTEGPQOTWIISEPEWNEIFKKQFDPNFYGR-KL 100

QY 133 HPIVKOLEGD---GLISLKGEMAHHRKIIISPTFMENIKLIVPVVLSYVDMVDKMSD 189  
 DB 101 RPIIGDPEDKDEVNIFSTQGRWKRLRLTSSPSFNNISLRKVRNSVQSGTEILWNIOK 160  
 QY 190 LSENGEVVDVYEWFPQILFEDVYSRTAFSSYED-----GRAVFLQAOQML---- 237  
 DB 161 VRKNEIDIMLIV--YQETLGVISRIALQSSSNMFKNPILPKQALFNQSMHVFLLINGI 218  
 QY 238 ---CAEAFQV-----FIPGYRFFPTGNLKSRDKKIRKSLIKLIERRONAIDGE 287  
 DB 219 FPLIAGVFRKMSKVLPAASFIPAKIFDL-----IEVAQARIDQAKDEI--K 264  
 QY 288 GECKEPAKDILGLMIQA-----KNVTQDIVECKSFEPAG 322  
 DB 265 GVEPGEF--ODFIDLFLDARVPDVKTLSEANEDPAKSSVVKINKELTDEIILAQCFFVL 322  
 QY 323 FAGKOTSNLLTWTTILSNHPEWQAKARDEVLRVCGSRDVPFKDHVVKLKTLSMILNES 382  
 DB 323 AAGPDTTALSLSYATVLAHPEIQKLOEVDRECPDEI--FPDHLSTKYLECVKMET 381  
 QY 383 LRLYP--PIVATTIRAKSDVKLGKIKIPGTELLIPTIAVHDOAIKGDVNEFPARFAD 441  
 DB 382 LRLYPLDTTANTRKCMRETTINGVNFDEGNITQVDWTLHNPRIENGVEDVEFKERMEN 441  
 QY 442 GVPRAAKHPVGFIPGIGVATCIGOMLAILQAKLTIAVNIQRTFLHAFYQ 493  
 DB 442 GACHLEHNSYIFPFGSGPRQICQMLAQMEQKILIAQLKKEYSFRTTQVQ 493

Search completed: May 19, 2004, 10:27:52  
 Job time : 23 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 19, 2004, 10:19:50 ; Search time 17 Seconds  
(without alignments)

1592.734 Million cell updates/sec

Title: US-09-992-901-2

Perfect score: 2730  
Sequence: 1 MEESSSWTFPKVTLVSLVIL.....LTPQHGAPITFRLLTHED 520

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_42:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1101	40.3	524	1 CP72 CATRO	Q05047 catharanthu
2	523	19.2	509	1 CP41 RAT	P08516 rattus norv
3	503	18.4	518	1 YRVA CAEEL	Q27519 caenorhabdi
4	495	18.1	518	1 YRVA CAEEL	Q27515 caenorhabdi
5	492	18.0	511	1 CP45 RABIT	P14579 oryctolagus
6	487.5	17.9	504	1 CP38 MOUSE	Q64459 mus musculu
7	487.5	17.9	506	1 CP44 RABIT	P10611 oryctolagus
8	482	17.7	537	1 CP56 RAT	P51871 rattus norv
9	481.5	17.6	504	1 CP41 MOUSE	Q09147 mus musculu
10	481.5	17.6	519	1 CP4Y HUMAN	Q02928 homo sapien
11	481	17.6	511	1 CP4B RAT	P15129 rattus norv
12	480	17.6	519	1 YRVA CAEEL	Q27520 caenorhabdi
13	479.5	17.5	520	1 CP3G MOUSE	Q64481 mus musculu
14	478	17.5	520	1 YRVA CAEEL	Q27514 caenorhabdi
15	477	17.5	511	1 CP4B HUMAN	P13584 homo sapien
16	477	17.5	534	1 YS24 CAEEL	Q09553 caenorhabdi
17	476.5	17.5	524	1 CPFB HUMAN	Q09b16 homo sapien
18	476	17.4	522	1 CPFA RAT	P51869 rattus norv
19	475.5	17.4	510	1 CP46 RABIT	P14580 oryctolagus
20	474.5	17.4	504	1 CP42 RAT	P20816 rattus norv
21	473	17.3	511	1 CP4T RABIT	P14581 oryctolagus
22	470	17.2	524	1 CP41 RAT	P33774 rattus norv
23	467.5	17.1	520	1 YRVA CAEEL	Q27513 caenorhabdi
24	466	17.1	524	1 CPFC HUMAN	Q09b62 homo sapien
25	465.5	17.1	520	1 YRVA CAEEL	Q27517 caenorhabdi
26	464	17.0	511	1 CP4B MOUSE	Q64462 mus musculu
27	464	17.0	520	1 CP4B HUMAN	P98187 homo sapien
28	463.5	17.0	508	1 CP48 RAT	P24464 rattus norv
29	463	17.0	507	1 CP43 RAT	P20417 rattus norv
30	459	16.8	478	1 CP43 DROME	Q09y40 drosophila
31	458	16.8	526	1 CP55 RAT	P51870 rattus norv
32	455	16.7	503	1 CP39 RAT	P51338 rattus norv
33	454.5	16.6	504	1 CP32 RAT	P05183 rattus norv

## ALIGNMENTS

RESULT 1	CP72 CATRO	STANDARD;	PRT;	524 AA.
AC	005047;			
DT	01-FEB-1994 (Rel. 28, Created)			
DT	01-FEB-1994 (Rel. 28, Last sequence update)			
DT	28-FEB-2003 (Rel. 41, Last annotation update)			
DE	Cytochrome P450 72A1 (EC 1.3.3.9) (CYPLXIII) (Secologanin synthase)			
DE	(SIS).			
GN	CYP72A1 OR CYP72 OR P450CR3.			
OS	Catharanthus roseus (Roei periwinkle) (Madagascar periwinkle).			Q070537 mesocricetu
OC	Catharanthus roseus (Rosaceae); Streptophyta; Embryophyta; Tracheophyta;			Q29496 ovis aries
OC	Eukaryota; Viridiplantae; Streptophyta; Eudicotyledons; Core eudicots; asterids;			Q97427 drosophila
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;			Q27518 caenorhabdi
OC	Lamiales; Gentianales; Apocynaceae; Rauvolfioideae; Vincaceae;			Q29981 blaberus di
OC	Catharanthus.			P51828 oryctolagus
NCBI_TaxID=4058;				P78329 homo sapien
[1]				Q97427 drosophila
SEQUENCE FROM N.A.				P04800 rattus norv
STRAIN=cv. CP3A;				Q27589 drosophila
Vetter H.-P., Mangold U., Schroeder G., Warner F.-J.,				Q64464 mus musculu
Meier-Reichardt D., Schroeder U.,				P47787 sus scrofa
"Molecular analysis and heterologous expression of an inducible				
cytochrome P-450 protein from periwinkle (Catharanthus roseus L.).";				
Plant Physiol. 100:998-1007(1992).				
[2]				
SEQUENCE OF 469-524 FROM N.A.				
STRAIN=cv. G. Don;				
MDLINE=93283641; PubMed=8507838;				
Meijer A.H., Souer E., Verpoorte R., Hoge J.H.C.,				
"Isolation of cytochrome P-450 cDNA clones from the higher plant				
Catharanthus roseus by a PCR strategy.";				
Plant Mol. Biol. 22:379-383(1993).				
[3]				
FUNCTION.				
STRAIN=cv. CP3A;				
MDLINE=20515722; PubMed=11135113;				
Immler S., Schroeder G., St-Pierre B., Crouch N.P., Hotze M.,				
Schmidt J., Strack D., Matern U., Schroeder J.,				
"Indole alkaloid biosynthesis in Catharanthus roseus: new enzyme				
activities and identification of cytochrome P450 CYP72A1 as				
secologanin synthase.";				
Plant J. 24:797-804(2000).				
-1- FUNCTION: Converts loganin into secologanin.				
-1- CATALYTIC ACTIVITY: Loganin + NADPH + O(2) = secologanin + NADP(+) + 2 H(2)O.				
-1- PATHWAY: Indole alkaloids biosynthesis.				
-1- SUBCELLULAR LOCATION: Vacuolar membrane (Probable).				
-1- TISSUE SPECIFICITY: Upper and lower leaf epidermis.				
-1- SIMILARITY: Belongs to the cytochrome P450 family.				
-----				
This SWISS-PROT entry is copyrighted. It is produced through a collaboration				
between the Swiss Institute of Bioinformatics and the EMBL outstation -				
the European Bioinformatics Institute. There are no restrictions on its				
use by non-profit institutions as long as its content is in no way				
modified and this statement is not removed. Usage by and for commercial				
entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a>				

or send an email to [license@lsb-sib.ch](mailto:license@lsb-sib.ch).

CC -----  
 CC EMBL: L10081; AAA33106.1; -;  
 DR EMBL: X69775; CAA49430.1; -;  
 DR PIR: S35168; S35168.  
 DR PIR: T09944; T09944.  
 DR HSSP: P14779; 1JPZ.  
 DR InterPro: IPR001128; Cytochrome\_P450.  
 DR Pfam: PF00067; P450.1.  
 DR PRINTS: PR00385; P450.  
 DR PROSITE: PS00086; CYTOCHROME\_P450; 1.  
 KW Oxidoreductase; Monooxygenase; Electron transport; Transmembrane;  
 KW NADP; Heme; Alkaloid metabolism.  
 FT TRANSMEM 12 32 POTENTIAL.  
 FT METAL 470 470 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).  
 FT VARIANT 190 190 I -> L.  
 FT VARIANT 194 194 Q -> E.  
 FT VARIANT 223 223 E -> D.  
 FT VARIANT 312 312 K -> R.  
 FT VARIANT 318 318 S -> T.  
 FT VARIANT 403 403 V -> I.  
 FT VARIANT 405 405 K -> E.  
 FT VARIANT 411 411 S -> P.  
 FT VARIANT 411 411 S -> P.  
 SQ SEQUENCE 524 AA; 60557 MW; EFSDB64E43C751B8 CRC64;

Query Match 40.3%; Score 1101; DB 1; Length 524;  
 Best Local Similarity 41.4%; Pred. No. 4e-63;  
 Matches 215; Conservative 112; Mismatches 166; Indels 26; Gaps 8;

QY 15 VLAVLSLVYKGSLL---WRPRKIEEHKSGKGRPRPHFFGVKVELVOMLKAS 71  
 DB 12 IAAITFALVWAMRVLDMAWFTPKRLKRLQGGFRGNRYRFLVGVDSGMMGEALS 71  
 QY 72 HPMPSHNLPRVLSFYHWRKIGATFLVWFPTFLVADPLIREIFSKSEFYKXN 130  
 DB 72 KPMENNDIVRLMHPHINTINTGRNSFTWNGRIPRIHWEDELKEVLTHSKYQKXF 131  
 QY 111 EAH-PLVKGEGDGLSLKGEKRAHNRKIISPTFMENLTLVPLVLSKSTVDWVDSK 189  
 DB 132 DVHNPVLFLL-TGVSEFGAKMKSHRIISPAITLKLKSMPPAICYHDLTWEXI 190  
 QY 190 LSENGEVVDVYEWFOITTEVDVSRITAFSGSYEDGRAVFLQAOQMLCAEAFQKFI 249  
 DB 191 AEKQSHVDLFPPTDVLTSLVSKVAFSGSYEGKIFLLKELMDLITDCRDVYI 250  
 QY 250 YRFPTGNLKRKLDEIKRSLIKLIERRQNAIDEGEGCEKPAKDLGLMIQAK-- 307  
 DB 251 WSYLPTKKNKMKKEINKEI-TDMRFLINKRMKAL-----KAGBPGEDDLGLVLSN 304  
 QY 308 -----NVTVODIVERCEKFFPAKQTTNLTWTTLILSMHPEWQAKRDEVR 356  
 DB 305 EIQKQKKGQSGMSINDVIECKLFYRAGQTTGLVLTITLILSKHPEWQAKRDEVR 364  
 QY 357 VCGSDVPTKDVVYKLTSLMILNESLRLPPIVATIRAKSDVYKGGYKIPGTELLP 416  
 DB 365 AFG-KNKPDEFRLNKLKYSMLILEVRLYPVLDLRIYKQTKGSIYIIPAGTOMLP 423  
 QY 417 IIAVHDAINGNDVNEFNPARPDGVPRAKHVGPIPLQIGRTIGNLLIQAOKT 476  
 DB 424 TWMHREKSLWGEDMEENRPFVDGVANAKKANVTYLPISWGRVLCGNPALLQAOKG 483  
 QY 477 LAVMIQRTFLAPTYQHAPVLMILVYPOHGAFTFERL 515  
 DB 484 LAMTLQRFKFDVAASYAAPFTILTVCQFGSHVYKKL 522

RESULT 2  
 CP41\_RAT STANDARD: PRT; 509 AA.  
 AC P08516;  
 DT 01-AUG-1988 (Rel. 08, Created)  
 DT 01-AUG-1988 (Rel. 08, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Cytochrome P450 4A1 (EC 1.14.15.3) (CYP1A1) (lauric acid omega-  
 hydroxylase) (P450-1A-omega 1) (P452).  
 GN CYP4A1 OR CYP4A-1.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_Taxid=10116;  
 [1]  
 RN SEQUENCE FROM N.A.  
 RP TISSUE=Liver;  
 RC MEDLINE=87109183; PubMed=3027069;  
 RX Hardwick J.P., Song B.-U., Huberman E., Gonzalez F.J.;  
 RA "Isolation, complementary DNA sequence, and regulation of rat hepatic  
 RT lauric acid omega-hydroxylase (cytochrome P-450A omega).  
 RT Identification of a new cytochrome P-450 gene family.";  
 RL J. Biol. Chem. 262:801-810(1987).  
 [2]  
 RN SEQUENCE FROM N.A.  
 RP STRAIN=Wistar; TISSUE=Liver;  
 RC MEDLINE=88312998; PubMed=3410047;  
 RX Barnshaw D., Dale J.W., Goldfarb P.S., Gibson G.G.;  
 RA "Differential splicing in the 3' non-coding region of rat cytochrome  
 RT P-452 (P450 IVA1) mRNA."  
 RL FEBS Lett. 236:357-361(1988).  
 [3]  
 RN SEQUENCE FROM N.A.  
 RP TISSUE=Liver;  
 RC MEDLINE=89356271; PubMed=2766932;  
 RX Kimura S., Hanloka N., Matsunaga E., Gonzalez F.J.;  
 RA "The rat clobazam-inducible CYP4A gene subfamily. I. Complete  
 RT intron and exon sequence of the CYP4A1 and CYP4A2 genes, unique exon  
 RT organization, and identification of a conserved 19-bp upstream  
 RT element.";  
 RL DNA 8:503-516(1989).  
 [4]  
 RN COVALENT HEME ATTACHMENT.  
 RP MEDLINE=21179192; PubMed=1119583;  
 RX Hoch U., Ortiz de Montellano P.R.;  
 RA "Covalently linked heme in cytochrome P450A fatty acid  
 RT hydroxylases.";  
 RL J. Biol. Chem. 276:11339-11346(2001).  
 [5]  
 RN COVALENT HEME ATTACHMENT, AND MUTAGENESIS OF GLU-320.  
 RP MEDLINE=21935389; PubMed=11821421;  
 RX Lebrun L.A., Hoch U., Ortiz de Montellano P.R.;  
 RA "Autocatalytic mechanism and consequences of covalent heme attachment  
 RT in the cytochrome P450A family.";  
 RL J. Biol. Chem. 277:12755-12761(2002).  
 CC -1- FUNCTION: Cytochromes P450 are a group of heme-thiolate  
 CC monooxygenases. In liver microsomes, this enzyme is involved in an  
 CC NADPH-dependent electron transport pathway. It oxidizes a variety  
 CC of structurally unrelated compounds, including steroids, fatty  
 CC acids, and xenobiotics.  
 CC -1- CATALYTIC ACTIVITY: Octane + reduced rubredoxin + O(2) = 1-octanol  
 CC + oxidized rubredoxin + H(2)O.  
 CC -1- SUBCELLULAR LOCATION: Membrane-bound. Endoplasmic reticulum.  
 CC -1- INDUCTION: By clofibrate.  
 CC -1- SIMILARITY: Belongs to the cytochrome P450 family.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.lsb-sib.ch/announce/>  
 CC or send an email to [license@lsb-sib.ch](mailto:license@lsb-sib.ch)).

DR Pfam; PF00067; P450; 1.  
 DR PRINTS; PR00385; P450.  
 DR PROSITE; PS00086; CYTOCHROME P450; 1.  
 KM Oxidoreductase; Monooxygenase; Electron transport; Membrane; Heme;  
 KM Microsome; Endoplasmic reticulum; HEME (COVALENT).  
 FT BINDING 320 320  
 FT METAL 456 456  
 FT MUTAGEN 320 320  
 FT COMFLICT 341 341  
 FT SEQUENCE 509 AA; 58214 MW; 11B5E102476A4DC2 CRC64;

Query Match 19.2%; Score 523; DB 1; Length 509;  
 Best Local Similarity 28.4%; Pred. No. 2,6e-26;  
 Matches 151; Conservative 105; Mismatches 218; Indels 58; Gaps 17;

QY SSSWFIKXVLSVLSLVKQWLSLWMPKRIEHPFKOGIRGPPYHFIQNV----- 59  
 DB SSGFLQVAVSLGLL-LLVKAVQFLQKQWLK--AFQGPSPFFHFWFGHQFGD 68  
 QY 60 KELVGMKASHPMPSENILPRVLSFYHMKRTIGATFLVWF-GPTFLVADPDILR 118  
 DB KELQOIMTCVENPSPAF-----PR-----WFGSKAYLLVDPDYMK 105  
 QY 119 EIFSKSEFEYKNEHNPVVKQLEGDGLSLKGEKVAHHRKISPTFMENIKLIVPVLS 178  
 DB 106 VILGRSD-PRANGVYRLAPWIGYGLLLNGQWPHGRKMLTPAFHYDLKRYVNMADS 164  
 QY 179 VTDWVDMKSKLSENGEVEVDVYEWFOILTEVISRTAF--GSYEDGRAVFLQAOQM 235  
 DB 165 IRLMLDKMEQLAGDSSIR--FQHSIMTLDTVMCAFSHNSVQVDNYSYIOAGN 222  
 QY 236 L--LCAPAFQKVPFPGVRF--FPTRGMLKSR--KLQKEIRKSLKLEHRRONADIGE 289  
 DB 223 LNDLFHSVRNIFQNDNTYFNSSNGELFNRAQCLADHTDGYIKRQQLONA--GELE 280  
 QY 290 ECKEPAKDLGLMIOAK-----NVTVDIVECKSFPPAGKOTSNLLTWTTLLSMHP 344  
 DB 281 KYKGRRLDPLDILLARMENGSLSKDLRAEVDTFMEGHDTTASGVSWIFVALATHP 340  
 QY 345 EWQAKADEVLRVCGSDVPTKDVVVKLTKLSMLNESLRLPYVATIRPAKSDVKL-G 403  
 DB 341 KHQORCRBEVQSVIGDSSITWDHLDQIPYTWICIKALRLYPPVGIVRELSTVTFPD 400  
 QY 404 GYKIPCGTELLIPFIAVHDOAIWGDVNEFNPNARFADGVPRAAKPPVFIFPGQVTC 463  
 DB 401 GRSIPKGIQVTLSTYGLHHPKWPNP-EVFDSPRFPDPSR---HSHFELPPSGARVC 456  
 QY 464 IGMNAILQAKLTLAVMIQRTFHLAPTYGHAFVLMLLYFGHAGITTRRL 515  
 DB 457 IGKQFAMSEMKVIALTLTLRFELLDPKVPILPRLVLKSKNGIYLYLKKL 508

## RESULT 3

IRVA\_CAEEL STANDARD; PRT; 518 AA.  
 ID IRVA\_CAEEL  
 AC Q27519;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Putative cytochrome P450 CYP13A7 (EC 1.14.-.-).  
 GN CYP13A7 OR T10B9.10.  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;  
 OC Rhabditidae; Peloderinae; Caenorhabditis.  
 OX NCBI\_Taxid=6239;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Bristol NZ;  
 RA Gardner A.;  
 RL submitted (MAR-1995) to the EMBL/GenBank/DBJ databases.

CC -!- FUNCTION: Cytochromes P450 are a group of heme-cholate  
 CC monooxygenases. They oxidize a variety of structurally unrelated  
 CC compounds, including steroids, fatty acids, and xenobiotics.

CC -!- SIMILARITY: Belongs to the cytochrome P450 family.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; Z48717; CAAB8609.1; -  
 CC PIR; T24783; T24783.  
 CC HSP; P14779; 1JPZ.  
 CC Womsep; T10B9.10; CE01655.  
 CC InterPro; IPR001128; Cytochrome\_P450.  
 CC Pfam; PF00067; P450; 1.  
 CC PRINTS; PR00385; P450.  
 CC PROSITE; PS00086; CYTOCHROME P450; 1.  
 CC Hypothetical protein; Oxidoreductase; Monooxygenase; Heme.  
 CC METAL 464 464  
 CC SEQUENCE 518 AA; 58999 MW; CC04283BF87B9EA7 CRC64;

Query Match 18.4%; Score 503; DB 1; Length 518;  
 Best Local Similarity 26.6%; Pred. No. 5e-25;  
 Matches 148; Conservative 109; Mismatches 180; Indels 120; Gaps 20;

QY SVLSLVIVYG-WSLMWRPKRIEHPFKOGIRGPPYHFIQNVKELVGMKASHPMP 75  
 DB 4 SIMLAIFGISTYLM-----IMSWIKYKRGRLGFLGVIRK----- 45  
 QY 76 FSHNLPRLVSFYHMKRTIGATFLVFGPPTFLVADPDILREIFSK--SEFY----- 127  
 DB 46 FTVYENPGLKLF-SEWTKKXGPGVYGTGEVETLVISDPEFHEVFGKQDFNFKULTA 104  
 QY 128 -----ENENHNPVVKQLEGDGLSLKGEKVAHHRKISPTFMENIKLIVPVLSYTM 182  
 DB 105 IQGDPNNKRPVPLV-----AAQGHWRKRLTLASTFENKSLRKMTGVEASVTEL 155  
 QY 183 VDKMSDKLSENGEVEVDVYEWFOILTEVISRTAFSS-----YEDGRAVFLQAO 233  
 DB 156 V-RSLEKASAEKGT-LDMLEYQEFMDIIGWAWQEKSLMFRNMLDKVATIFREGAN 213  
 QY 234 QMLICAPAFQKVPFPGVRFPTRGMLKSR-KSLDKETKSLKLEHRRON--AIDGEE 289  
 DB 214 NVFMISGIFPFVIAIRNIFAKPRLQMAVDIOSILKALNRLQREADEFAGLEPSGE 273  
 QY 290 ECKEPAKDLGLMIOA-----KNVTVDIVECKSFPPAG 325  
 DB 274 -----PQDIDFLQARSVPDFBGAQDPAKSEVLKVDHLPFDEILIGLFFLLAG 327  
 QY 326 KQTTSNLLTWTTLLSNHPWQAKADEVLRVCGSDVPTKDVVVKLTKLSMLNESLRL 385  
 DB 328 YDTALSLSYSLATHPETIQKQOEVDRECPDEV-TFDQLSLKTLKCVVEBALRL 386  
 QY 386 YPPIVATL---RAKSNVVLGKYGKIPCGTELLIPFIAVHDOAIWGDVNEFNPNAR 438  
 DB 387 YP--LASTLVNRKCLKTLNVGLMEIENAGVINVDTSIHHDKRVQGDVNEKPRKWSG 444  
 QY 439 ----FADGVPRAAKHPVGF-PFGIVATCIQGNVAILQAKLTLAVMIQRTFHLAP 490  
 DB 445 DELFFAKG-----GYLPFGMGPRICIGMLAMEMKMLTNILKNYTFETTPETVI 495  
 QY 491 -----TYQNAPTVLM 501  
 DB 496 PLKLVGATTAAPSSVLL 512

## RESULT 4

IRV3\_CAEEL STANDARD; PRT; 518 AA.  
 ID IRV3\_CAEEL  
 AC Q27515;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)

28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Putative cytochrome P450 CYP1A6 (EC 1.14.-.-).  
 GN CYP1A6 OR T10B9.3.  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidae;  
 OC Rhabditidae; Peloderidae; Caenorhabditis.  
 OX NCBI\_TaxID=6239;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Bristol N2;  
 RA Gardner A.;  
 RL Submitted (MAR-1995) to the EMBL/GenBank/DBJ databases.  
 CC -1- FUNCTION: Cytochromes P450 are a variety of structurally unrelated  
 CC monooxygenases. They oxidize a variety of structurally unrelated  
 CC compounds, including steroids, fatty acids, and xenobiotics.  
 CC -1- SIMILARITY: Belongs to the cytochrome P450 family.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; 248717; CA88605.1; -.  
 DR F01; T24779; T24779.  
 DR HSSP; P14779; J1P2.  
 DR WormRep; T10B9.3; CE01657.  
 DR InterPro; IPR001128; Cytochrome\_P450.  
 DR Pfam; PF00067; P450.1.  
 DR PRINTS; PR00385; P450.  
 DR PROSITE; PS00086; CYTOCHROME\_P450.1. Monooxygenase; Heme.  
 DR Hypothetical protein; Oxidoreductase (By SIMILARITY).  
 KM Hypothetical protein; Oxidoreductase (By SIMILARITY).  
 FT METAL 463 463 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).  
 SQ SEQUENCE 518 AA; 59444 MW; 381787A30E2792CA CRC64;  
 Query Match 18.1%; Score 495; DB 1; Length 518;  
 Best Local Similarity 26.7%; Pred. No. 1.6e-24;  
 Matches 142; Conservative 101; Mismatches 197; Indels 92; Gaps 17;  
 QY 15 VLSVLSVYKQKSLMWRPKRBEHFKGIGPPHYFTGVKSLVGMKASHRM 74  
 DB 1 MIFVLISAVLIGVTVSW- - - - - IWSYFIRKIGKGP- RGFPG- - - - - IGMILQTDHEN 49  
 QY 75 PSHNILLPRVLSFYHMKRYGATPLVWFGPTFLTVADPLDRLREIFSK- - - SEFYKNEA 132  
 DB 50 P- - - - - PFLKIDMTKQYQPVYGFEGEQGMIISEPMVNEIRKQFDNFYGR- KL 100  
 QY 133 HPLVQLQEGD- - - GLSLIKGEKMAHRIKIIPTFHENMLKLVPLVSKVTDWYKMSK 189  
 DB 101 RLISGPEKQKRVNIFSTQGRKWKRLFTLSPPSSNNLSRKRVNSVCEGTEIIMNTEOK 160  
 QY 190 LSENGSEVDYVEMQOILTEDVTSRTAFSSYED- - - - - GRAVFRLOAQOMLL- - - 237  
 DB 161 VRKNEDIDLVIV- - - YQVETLVGISRIALQGSSESNFKNPRLPKYQALFNSSMHVFLITGI 218  
 QY 238 - - - - - CAEAFQKV- - - - - FIPGYRFPFRGNLKSRLDKLIERKSLKLIERRRONAIDG 287  
 DB 219 PPLAGVFRKSKMLPASFIPIAFKIFDL- - - - - LEVAVOARIQORADEI- - - K 264  
 QY 288 GEECEPRAKDLIGMIQ- - - - - KNTVQDIYBECSFF 332  
 DB 265 GVEPGEF- - - QDFIDFLDARVVDVKILSGANEDPAKSVYKINKELTFDEILIQCVFL 332  
 QY 323 FAGKQTSNLLTWITLLSMHEWQAKARDEVLRVCGRDVPTQDVHVKLTLSMILNES 382  
 DB 323 AAGFTTALSVAATVLATHPDIQTLQAEVDRCPDEI- - - PFDHLSKLKYLECWAKET 381  
 QY 383 LALYP- - - PIYATIRAKSDYKLGKYGKIPGSTELLIIIVHNDQAIWGDVNEPAPFAD 441  
 DB 382 LRLYPGLTANTRKCKRETTINGVNFDBSGNIQVDTWTLHHNPRIMGEDVDFEERWEN 441

QY 442 GVRPAANPVGFIPIFGLGVFTCTGONLALLOAKTLTANMIOREFTHLAPTYQ 493  
 DB 442 GACEHLEHNSYIPFGSGPQICIGRLAQMEOKILLAQILKEYSFTTKNTQ 493  
 RESULT 5  
 ID CP45\_RABIT STANDARD; PRT; 511 AA.  
 AC CP45\_RABIT  
 DT 01-JAN-1990 (Rel. 13, Created)  
 DT 01-JAN-1990 (Rel. 13, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Cytochrome P450 4A5 precursor (EC 1.14.15.3) (CYP1A5) (Lauric acid  
 DE omega-hydroxylase).  
 GN CYP4A5  
 OS Oryctolagus cuniculus (Rabbit).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.  
 OX NCBI\_TaxID=9986;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Kidney;  
 RX MEDLINE=90254128; PubMed=2340280;  
 RA Johnson E.F., Walker D.L., Griffin K.J., Clark J.E., Okita R.T.,  
 RA Neuhoff A.S., Masters B.S.S.;  
 RT "Cloning and expression of three rabbit kidney cDNAs encoding lauric  
 RT acid omega-hydroxylase.";  
 RL Biochemistry 29:873-879 (1990).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Kidney;  
 RX MEDLINE=91192021; PubMed=2013275;  
 RA Yokotani N., Kusunose E., Sogawa K., Kawashima H., Kinoshita M.,  
 RA Kusunose M., Fujii-Kuriyama Y.;  
 RT "cDNA cloning and expression of the mRNA for cytochrome P-450kd which  
 RT shows a fatty acid omega-hydroxylating activity.";  
 RL Eur. J. Biochem. 196:531-536 (1991).  
 CC -1- FUNCTION: Cytochromes P450 are a group of heme-chlorate  
 CC monooxygenases. In liver microsomes, this enzyme is involved in an  
 CC NADPH-dependent electron transport pathway. It oxidizes a variety  
 CC of structurally unrelated compounds, including steroids, fatty  
 CC acids, and xenobiotics.  
 CC -1- CATALYTIC ACTIVITY: Octane + reduced rubredoxin + O(2) = 1-octanol  
 CC + oxidized rubredoxin + H(2)O.  
 CC -1- SUBCELLULAR LOCATION: Membrane-bound. Endoplasmic reticulum.  
 CC -1- INDUCTION: P450 can be induced to high levels in liver and other  
 CC tissues by various foreign compounds, including drugs, pesticides,  
 CC and carcinogens.  
 CC -1- SIMILARITY: Belongs to the cytochrome P450 family.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; M28655; AA31329.1; -.  
 DR EMBL; X57209; CAA40493.1; -.  
 DR PIR; A34260; A34260.  
 DR HSSP; P14779; J1P2.  
 DR InterPro; IPR001128; Cytochrome\_P450.  
 DR Pfam; PF00067; P450.1.  
 DR PRINTS; PR00385; P450.  
 DR PROSITE; PS00086; CYTOCHROME\_P450.1.  
 DR Oxidoreductase; Monooxygenase; Electron transport; Membrane; Heme;  
 KW Microsome; Endoplasmic reticulum.  
 FT PROPEP 1  
 FT CHAIN 5 511 CYTOCHROME P450 4A5.  
 FT METAL 458 458 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).  
 FT CONFLICT 435 435 G -> S (IN REF. 2).  
 FT CONFLICT 477 477 V -> L (IN REF. 2).

SQ SEQUENCE 511 AA; 58357 MW; 11D174BFC8B2A268 CRC64;  
 Query Match 18.0%; Score 492; DB 1; Length 511;  
 Best Local Similarity 27.7%; Pred. No. 2.5e-24;  
 Matches 146; Conservative 103; Mismatches 219; Indels 60; Gaps 17;

QY 16 LSVLSLVIVKGMIL-----MWRPKIEHFSSQSGRGPYHFIQNVLELTVG 64  
 DB 15 LSGLLQVAAALGILLIKAAQLYLRQWMLRALQ-----PCPCPFHMLHSHB--- 65  
 QY 65 MMLKASHPMPFESH-ILPVLSPFYHMKRYGATFLVWEGPT-FRLTVADPDLREIFS 122  
 DB 66 -----FQNNQLQQLTK-----WVEKFPACPHWIGKVKVRQVLDPPYMKVILG 110  
 QY 123 KSEPEKKEAHPLVYQLEGGDGLSLKGEKNAHKKIISPTFHEMLKLAVPVYLKSVTDM 182  
 DB 111 RSD-PKSRGSYTFVAPWIGYGLLLNQPMFOHRRMLTPAFHYDILKPYVGLWDSVQIM 169  
 QY 183 VDKMEDKLSNGEVEVDVYEWFOILTEDVISRTAF---GSSYEDGR-AVERLQAOQML-- 236  
 DB 170 LDKMQLVSGDSLE--VFQDISLMTLDTIKKAFSYQGSVQDSRMSQSYIQAVGDLNN 227  
 QY 237 LCAEFQKVFIPG---YFPPTKGNLKRKLDKRIKSLKLERRQNAIDEGECCKE 293  
 DB 228 LVFAFVRNIFHQSDTIYRLSP-EGRLSHRACQLAHEHTDRVIOQRKALQOESELEKVR 286  
 QY 224 PAKKLLGLMTQAK-----NTVQDIYECKSPFPFAGKQTSNLTWTTLISMHPBWOA 348  
 DB 287 KERLDFLVLLPARKENGSSLSDDQLRAVDVTFEFGHDTTASVSMIFALATHENHGH 346  
 QY 349 KARDEVLRVCGSRDVPIDHVKLKTLSMINSRLRPPVATIRPAKSDVKL-GGYKI 407  
 DB 347 RCRERIGLGDGASITWEHLDPMYTMCIKEMARLYPPPAISRLSSPVPFPDGRSL 406  
 QY 408 PGCTLLPIIAVHDDQAIWGDVNERPAPFACGVPRAAKHPGCFPLGAVRTICIGN 467  
 DB 407 PGKFTVTLISYGLHNPNVWPP- EVFDPGRF---TOSASHNAFLPFGGARNCIGKQ 462  
 QY 468 LAILOAKLTAVMIQRTFHLAPTYQHAPVLYMLLYPOHGAPITFRLL 515  
 DB 463 FAMELKVAVALLTVRFELDPFRIRPRTARLVLSKNSNGIHLRLKL 510

RESULT 6  
 CP3B MOUSE STANDARD; PRT; 504 AA.  
 ID CP3B MOUSE  
 AC 06459;  
 DT 15-UTL-1999 (Rel. 38, Created)  
 DT 15-UTL-1999 (Rel. 38, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Cytochrome P450 3A11 (EC 1.14.14.1) (CYP11A1) (P-45011A1) (P-450UT)  
 GN CYP3A11 OR CYP3A-11.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=ddy; TISSUE=Liver;  
 RX MEDLINE=92223116; PubMed=1339292;  
 RA Yanagimoto T., Itoh S., Muller-Brosch D., Kamataki T.;  
 RT "Mouse liver cytochrome P-450 (P-45011A1): its cDNA cloning and  
 RL induction by dexamethasone." Acta 1130:329-332 (1992).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Liver;  
 RX MEDLINE=22388257; PubMed=12477932;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Scheef C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,

RA Diachenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Udell T.B., Toshiyuki S., Carnant P., Prange C.,  
 RA Rana S.S., Loguelfano N.A., Peters G.J., Abramson R.D., Mollath S.J.,  
 RA Bosak S.A., McEwan P.O., McKernan K.J., Malek U.A., Gay L.J., Hulyk S.W.,  
 RA Villalón D.K., Muzny K.C., Hale S., Garcia A.M., Gay L.J., Gibbs R.A.,  
 RA Fahy J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,  
 RA Blakesley R.M., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
 RA Butlerfield Y.S.N., Krzywicki M.I., Skalska U., Smalus D.E.,  
 RA Scherch A., Schein J.E., Jones S.U.M., Maria W.A.,  
 RT "Generation and initial analysis of more than 15,000 full-length  
 RL human and mouse cDNA sequences." Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).  
 CC - FUNCTION: Catalyzes erythromycin N-demethylation,  
 CC oxidation and testosterone 6 beta-hydroxylation.  
 CC - CATALYTIC ACTIVITY: NADPH + reduced flavoprotein + O(2) = NADP +  
 CC oxidized flavoprotein + H(2)O.  
 CC - SUBCELLULAR LOCATION: Membrane-bound. Endoplasmic reticulum.  
 CC - TISSUE SPECIFICITY: Highly expressed in liver.  
 CC - INDUCTION: By dexamethasone.  
 CC - SIMILARITY: Belongs to the cytochrome P450 family.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; X60452; CA942981.1; -  
 CC EMBL; BC010528; AAH10528.1; -  
 CC PIR; S22334; A60564.  
 CC HSP; P14779; 1PBZ.  
 CC MGD; MG1:86609; CYP3A11.  
 CC InterPro; IPR001128; Cytochrome P450.  
 CC InterPro; IPR008072; EP450\_CYP3A.  
 CC Pfam; PF00067; P450.1.  
 CC PRINTS; PR01689; EP450IICYP3A.  
 CC PRINTS; PR00385; P450.  
 CC PROSITE; PS00086; CYTOCHROME P450, 1.  
 CC Oxidoreductase; Monooxygenase; Electron transport; Membrane; Heme;  
 CC Microsome; Endoplasmic reticulum.  
 CC METAL 443 443 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).  
 SQ SEQUENCE 504 AA; 57854 MW; E369AF71CE23F180 CRC64;

Query Match 17.9%; Score 487.5; DB 1; Length 504;  
 Best Local Similarity 26.1%; Pred. No. 4.7e-24;  
 Matches 146; Conservative 93; Mismatches 178; Indels 103; Gaps 17;

QY 13 VLVSVLSLVIVKGMILMMRPKRIEHEFKOSIRGPPYHFGVAKELVGMMLKASH 72  
 DB 4 VSLSLTGVVLAISLVLYRYGRKHELFKQGIIPG-----K 42  
 QY 73 PMPFSHNILPRVLSFYH-NR-----KIYATFLVWFGPTFLTVADPDLREIFSK-- 123  
 DB 43 PLPFL---LGYLVANVYKGLMKFMECYKKYKGTGFLPDGCPPLAVADPFIKXVLYKEC 98  
 QY 124 -SEPEKKEAHPLVYQLEGGDGLS-----LKGEMAHHRKISPTFHEMLKLAVPVYLK 177  
 DB 99 FGVFTNRDPFV-----GIMSVAISISDDMKRFRALSLPFLSGKLKEMFVI-E 150  
 QY 178 SYTDWVDKMSDLSNGEVEVDVYEWFOILTEDVISRTAFSS-----YEDGRA 226  
 DB 151 QYGDILVYKLRQAKGK-PYTMKVLYGAYMDVITSPGVNDSLNPPDPVEYAKK 209  
 QY 227 VERLQAOQMLCAEFQKVFIPGR-----FPPTRGNLKRKLDKRIKSLKLERRRQ 261  
 DB 210 LRFDFPDLFLSVALFPFLTPVYEMLNICMFP-----KDSIEFFKKFVDYKKE 258



```

QY 282 NAIDGEGECKEPPAAKDLGLMTIOAKN-----VTVQDIVECKSFPPAK 326
D 259 SRLDS-----KOKHRDPLQLMNNSHNSKDVSHKALSDWEITAGOSII-----FIRAGY 308
QY 327 QTSNLTMTTLLTSLWHPWOKARDEVLKRVGSGSDVPTKHVYKLTLSMLNESRLY 386
D 309 ETTSTSLTSLHSLATHPDIOCKLQDEIDALPKAPPTTPTVMEETLDMVNLNETRLY 368
QY 387 PPIVATIRAKSDVYKGYKIPCGTELLIPITIAVHDOAIWGNVNEFPAPADGVPA 446
D 369 PIANRLERVCCKDVELNGVYIPKGSVTMIPSYALHDPQHW-SEPEEFQPERFSK-ENKG 426
QY 447 AAGPVGFIPEGCGVTCIGONLAIIOAKLTLMVIOREPT 486
D 427 SIDPYVLLPFGNGPNCLOMRPALMMKLTATKIMQNSF 466

RESULT 7
CP44_RABIT STANDARD; PRT; 506 AA.
ID ID_CP44_RABIT STANDARD; PRT; 506 AA.
AC P1061;
DT 01-JUL-1989 (Rel. 11, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Cytochrome P450 4A4 (EC 1.14.14.1) (CYP1A4) (Prostaglandin omega-
hydroxylase) (P450-P-2).
GN CYP4A4.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_Taxid=9986;
RN [1]
RP SEQUENCE OF 17-506 FROM N.A., AND SEQUENCE OF 1-25.
RX MEDLINE=88007548; PubMed=365614;
RA Matsubara S., Yamamoto S., Sogawa K., Yokotani N., Fujii-Kuriyama Y.,
RA Hanu M., Shively J.E., Gotch O., Kusunose E., Kusunose M.,
RT "cDNA cloning and inducible expression during pregnancy of the mRNA
RT for rabbit pulmonary prostaglandin omega-hydroxylase (cytochrome
RT P-450P-2).";
RI P. Biol. Chem. 262:13366-13371 (1987).
CC -1- FUNCTION: Cytochromes P450 are a group of heme-thiolate
CC monooxygenases. In liver microsomes, this enzyme is involved in an
CC NADPH-dependent electron transport pathway. It oxidizes a variety
CC of structurally unrelated compounds, including steroids, fatty
CC acids, and xenobiotics.
CC -1- CATALYTIC ACTIVITY: RH + reduced flavoprotein + O(2) = ROH +
CC oxidized flavoprotein + H(2)O.
CC -1- SUBCELLULAR LOCATION: Membrane-bound. Endoplasmic reticulum.
CC -1- INDUCTION: P450 can be induced to high levels in liver and other
CC tissues by various foreign compounds, including drugs, pesticides,
CC and carcinogens.
CC -1- SIMILARITY: Belongs to the cytochrome P450 family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: J02818; AAA31232.1; -.
CC PIR: S32315; A29368.
CC HSSP: P14779; IUPZ.
CC InterPro: IPR001128; Cytochrome_P450.
CC Pfam: PF00067; P450.1.
CC PRINTS: PR00385; P450.
CC PROSITE: PS00086; CYTOCHROME_P450; 1.
CC Oxidoreductase; Monooxygenase; Electron transport; Membrane; Heme;
CC Microsome; Endoplasmic reticulum.
CC METRAL 453 453 ION (HEME AXIAL LIGAND).
CC FT SEQUENCE 506 AA; 58525 MW; 50CB041463B9B50 CRC64;
CC SQ

```

```

Query Match 17.9% Score 487.5; DB 1; Length 506;
Best Local Similarity 27.2%; Pred. No. 4,78-24;
Matches 143; Conservative 105; Mismatches 221; Indels 57; Gaps 14;

QY 16 LSVILSLVIVXGSLMWRPKRIEY-----FSKQIGRPYFFIGYKELVGMMLKAS 70
D 11 LSGILQVAAALGILLILLKAAQVLYHROWLLBALQCPFPFLLGHSRE-----61
QY 71 SHMPFSNNILPRVLSFHHKRYGATFLVWF-GPFFRLTVADPDQIREIFSKSEFEYK 129
D 62 -----FQND---QELRIQKVEKFPQACPMWSLGNKARLLVDDPDYKILGRSD-PKA 112
QY 130 NEAHPVLVQLBDDGLSLKGEKMAHHRKISPTFHEMLKLLVPLVKSVTDVWYKXSDK 189
D 113 PANYLMTMPWGYGLLLLDGQTFQHRMLTPAHVYDLKRYGLAMDSVQIMLDPEQL 172
QY 190 LSENGEVDVYEMFQILTEDVIGRTAF---GSSYEGRAVFRLOAQOML--LCAEAFQK 244
D 173 ISQDSLE-IFQHSMTLDTIMKCAFSYQSVQDRNSHYIQALNDLNNLVFYARN 230
QY 245 VF-----IPGYRFPPTRGNTKSRKLDKEIRKSLKLTERRRONALDGECKEPA 235
D 231 VHQSDFLYRLSPBGRLLFRAOQLAHENTDVIQ-----QRAQQLQEGELEKVRKR 283
QY 296 ADLLGLMTIOAK-----NTVQDIVECKSFPPGKQTSNLTMTTSLMHPWQAKA 350
D 284 RLDPLDLVLFAMNGSSLSDQDLAEVDITMFGHDTAGSVWIFVYALATHEHQHC 343
QY 351 RDEVLKRVGSRVPRKDVVYKLTLSMLNESRLYPIVATIRAKSDVYL-OGYKIPC 409
D 344 RESIQGLDGDASITWEHLDQWPTTMCIKALKLIPVPSVTQLSKPVTFPGRSLPK 403
QY 410 GTELLIPITIAVHDOAIWGNVNEFPAPADGVPAKAPVGTTPFLGRTIGQNTLA 469
D 404 GVILFLSIYGLAHYNKWNQNP-EVDPDPFPA---PDSAYSHSHALPFGSARNCIGQFA 459
QY 470 IIOAKLTLMVIOREPTFLHAPYQAPVLMVLYPOHAPPTFRL 515
D 460 MELKVAVALTVRFELLDPDTRIPPIARVVLKSKGNIHLRLKL 505

RESULT 8
CP46_RAT STANDARD; PRT; 537 AA.
ID ID_CP46_RAT STANDARD; PRT; 537 AA.
AC P51871;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Cytochrome P450 4P6 (EC 1.14.14.1) (CYP1P6).
GN CYP4P6.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
OX NCBI_Taxid=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=Sprague-Dawley; TISSUE=Brain;
RX MEDLINE=96125358; PubMed=8554568;
RA Kawashima H., Strobel H.W.;
RT "cDNA cloning of three new forms of rat brain cytochrome P450
RT belonging to the CYP4P subfamily.";
RL Biochem. Biophys. Res. Commun. 217:1137-1144 (1995).
CC -1- CATALYTIC ACTIVITY: RH + reduced flavoprotein + O(2) = ROH +
CC oxidized flavoprotein + H(2)O.
CC -1- SUBCELLULAR LOCATION: Membrane-bound. Endoplasmic reticulum.
CC -1- TISSUE SPECIFICITY: High expression in liver and kidney. Lower
CC expression in brain.
CC -1- SIMILARITY: Belongs to the cytochrome P450 family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way

```

CC	modifies and this statement is not removed. Usage by and for commercial
CC	entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a>
CC	or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).
CC	-----
DR	EMBL; U93208; AAC52360.1; -.
DR	PIR; JC4534; JC4534.
DR	HSSP; P14779; IJFZ.
DR	InterPro: IP001128; Cytochrome_P450.
DR	Pfam: PF00067; p450_1.
DR	PRINTS; PR00385; P450.
DR	PROSITE; PS00086; CYTOCHROME_P450_1.
KM	Oxidoreductase, Monooxygenase; Electron transport; Membrane; Heme;
KW	Microsome; Endoplasmic reticulum.
FT	METAL 468 468 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).
SQ	SEQUENCE 537 AA; 61541 MW; 406D6761A2BEA7B9 CRC64;
Query Match 17.7%; Score 482; DB 1; Length 537;	
Best Local Similarity 30.2%; Pred. No. 1,16-23;	
Matches 128; Conservative 82; Mismatches 192; Indels 32; Gaps 14	
DQ	85 LSFFHWRIKYGATFVWFGEPTFR-LTVADPDLLRFISKFSEFEYKN--AHPLVKOLEG 141
DQ	74 MQFIHALGNFPDIHLSNVGPVPYPLRLVHNVIAPILQSAVAAPKENTLVGLKPMVG 133
DQ	142 DGLLSKEGKAHRRKIISPTFHENKLIVPVVLKSVTMDVRMSDKLSENGEVEVDY 201
DQ	134 DGLMSGAGKKMHHRLLTPAFHFPIILSKYIKNKSVNTMAKM-QRLTKGSARLDVF 192
DQ	202 EMNQILLTEVISRTAFG-----SYSDGRAVRLQAQQMLCAEAFQKVFIPGRFPP 254
DQ	193 EHISLTLLDSLCIFSPSPSCQSNSYTAIIIEISS---LIWKORGFLLDPLYTL 249
DQ	255 TRGNLKSRLLDEIRKSLLKIERRNAIDEG-ECKEPAAK---DLGLMTQAKN- 308
DQ	250 TADGRFRKA-CDVANFTDAVIFERRSTLNTQGVDSFLKRAATKTLLFDVLLAKDE 308
DQ	309 ----VVQGIIVECKSFPAKGQTTSNLITWTTLTSLMHPEMAKKADEVLRVCGSRDP 364
DQ	309 HGKGLSDVDIIRAEDTFMGGHDTYASLSWILYNLARHPYQRCQEVREILRRP-P 367
DQ	365 TK---DHVVKLTLSMLINLESRLTPVTATIRAKSDVKL-GGYKIPGSTELLPIAY 420
DQ	368 EELIEWDDLAQLPLTMCIESRLRPVTLISRCSODIYLPGRAVIPKNICVSIIFGV 427
DQ	421 HHDAQIWGDVNEFNARPADGVPPAAAGHPGFIPIFGIGVRTCTGGOTALQAKLTIAM 480
DQ	428 HHNPVSVM-DDPEVYNFRFDPENPO-KSPPLAFIPFSAGPRNCIGTFAMSEIKVALTT 485
DQ	481 IQRF 484
DQ	:: ::
DQ	486 LTRF 489
RESULT 9	
ID	C341_MOUSE STANDARD; PRT; 504 AA.
AC	O9JMA';
DT	28-FEB-2003 (Rel. 41, Created)
DT	28-FEB-2003 (Rel. 41, Last sequence update)
DT	28-FEB-2003 (Rel. 41, Last annotation update)
DE	Cytochrome P450 3A41 (EC 1.14.14.1).
GN	CYP3A41.
OS	Mus musculus (Mouse).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murine; Mus.
OX	NCBI_TaxID=10090;
RN	[1]
RP	SEQUENCE FROM N.A.
RC	STRAIN=ddy; TISSUE=Liver;
FX	MEDLINE=20239668; Pubmed=10775455;
RA	Sakuma T., Takei M., Endo Y., Kuroiwa W., Ohara A., Jankamjorn K.,
RT	Hotima R., Nemoto K.;
RT	"A novel female-specific member of the CYP3A gene subfamily in the

mouse liver." ;  
 Rt Arch. Biochem. Biophys. 377:153-162(2000).  
 CC -/- CATALYTIC ACTIVITY: RH + reduced flavoprotein + O(2) = ROH + oxidized flavoprotein + H(2)O.  
 CC -/- SUBCELLULAR LOCATION: Membrane-bound. Endoplasmic reticulum (potential).  
 CC -/- TISSUE SPECIFICITY: Expressed in liver. Also expressed in the kidneys of female mice, with traces in the stomach, ovary, and heart of female mice and in the testis of male mice.  
 CC -/- DEVELOPMENTAL STAGE: Detected immediately after birth in the livers of animals of both sexes, but increased with age in females; whereas it was gradually reduced in males, resulting in predominantly female-specific expression in livers.  
 CC -/- SIMILARITY: Belongs to the cytochrome P450 family.  
 ---  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 ---  
 DR EMBL; AB033414; BAA95951.1; -.  
 DR MGD; MG1:1658451; CYP3A4.1.  
 DR InterPro; IPR001128; Cytochrome\_P450.  
 DR InterPro; IPR008072; EP450\_CYP3A.  
 DR Pfam; PF00067; P450; 1.  
 DR PRINTS; PR01689; BP450IICYP3A.  
 DR PRINTS; PR00385; P450.  
 DR PROSITE; PS00086; CYTOCHROME\_P450; 1.  
 DR Oxidoreductase; Monooxygenase; Electron transport; Membrane; Heme; Microsome; Endoplasmic reticulum.  
 KM METAL. 443 443 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).  
 SQ SEQUENCE 504 AA; 57959 MW; 2BCC645BBE9CD46A CRC64;  
 Query Match 17.6%; Score 481.5; DB 1; Length 504;  
 Best Local Similarity 28.4%; Pred. No. 1, Le-23;  
 Matches 141; Conservative 89; Mismatches 170; Indels 111; Gaps 19;

18 VILSLVIYKGMGLMMRPKKIEBHFSKQGIKGPPIYHFIGNVAKELGVMKLKXSHRMPSS 77  
 :|:::|:  
 Db 13 VLIAITLV-----LTVRYGTTHGLFKQGSLPGP-----PLPF--46  
 :|:::|:  
 Qy 78 HTILPVLSFYHH-NR-----KIYGATFLWFGPTIRLTIVADPDILIREIFSK--SEFY 127  
 :|:::|:  
 Db 47 ---LGVLVLYIKGLMKFMECEKEKKQTGWISFDQMPLFIPTDPMIKAVLVYECPSVPT 103  
 :|:::|:  
 Qy 128 EKNEAHPLVKOLEGGDLIS-----LKGEKAHNRKIISPFFHNENLKTLPVPVLSKVTEM 182  
 :|:::|:  
 Db 104 NREEPFPV-----GIMSKAISISKDEWRKYALLSPFTSGKLKEMFPVL-EQYDI 155  
 :|:::|:  
 Qy 183 VDRKSOKLSENBE--VDVYVEWFQILTIEDVYSTAGGS-----YEDGRAVER 229  
 :|:::|:  
 Db 156 LVKITLMQBAEKKPVTMDVLGAYSI---DVISTSGVAWDLSLNPNDEPFYEKAGILR 212  
 :|:::|:  
 Qy 230 LOAQOMLLCAEAFOKVFIPGYR-----FFPRGNLKRKLDKEIRKSLKTLIRRRQNMI 284  
 :|:::|:  
 Db 213 VDPDFDELVSFVLPFLPYGVEMINICMFP-----KOSIEFFKFVVRMESRL 261  
 :|:::|:  
 Qy 285 DGESEBECKPAADLLGLMIQKN-----TYVDIVYECSKPPFAKQRT 329  
 :|:::|:  
 Db 262 DS-----KQGRVDFQLQMNNAHNSKSDKHSLSMETTAGSIV-----FFFAEYTT 311  
 :|:::|:  
 Qy 330 SNLTITWTITLISHNPEWQAKARDEVLRVCGRSDVPTRDYVVKTKTSLMINESBLRYPPI 389  
 :|:::|:  
 Db 312 SSTLSFLYLCLATHPIDIQCKQEIDETLTPNKAPPTDYTWEMELYDMVNETLRYPPIG 371  
 :|:::|:  
 Qy 390 VATIRRAKSVYKLGRIKPCGTELLIPLIIVHQDAIMGNDVNEFNPARPADGVPAACH 449  
 :|:::|:  
 Db 372 NRLERFCKQVDELNGVYIRPKGSTWMTISVALHPDHQH-PPEEFQERFSK-ENKGSID 429  
 :|:::|:  
 Qy 450 PVGIPRGGLGVRTIGQNLAIQAULTLAVMIORETTF 486



QY 363 VPTKHVVKIKTISMLNESRLYPVATITRAKSDVKL-CGYKIPCGTELLPIIAVH 421  
DB 360 SIYMHDDQPYTTCIKELRLYPVPGIGRELSPTFPDGRSLPGGINVLISYGLH 419  
QY 422 HDQAIWQNDVNEFNPAFADGVPRAKHVPGFIPGLGVRTICIGONLAILQAKTLAVMI 481  
DB 420 HNRKWPMPN-EVYDPRFRA---PGSKSHALPFGSGSRNCIGQFAMNELKATALT 475  
QY 482 QRTFHLAPTYQHAPVLMILYPOGAPITFRRLTN 517  
DB 476 LRFELPDPFRIPRIPIARVLTKSKNGIHRLRLRFP 511  
RESULT 11  
CP4B RAT STANDARD; PRT; 511 AA.  
ID CP4B RAT STANDARD; PRT; 511 AA.  
AC P5129;  
DT 01-APR-1990 (Rel. 14, Created)  
DT 01-APR-1990 (Rel. 14, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DB CYCROCHROME P450 4B1 (EC 1.14.14.1) (CYPR1B1) (P450-isozyme 5).  
GN CYP4B1 OR CYP4B-1.  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
OX NCBI\_TaxID=10116;  
RN [1]  
RX SEQUENCE FROM N.A.  
MEDLINE=89261667; PubMed=2725471;  
RA Gasser R., Philipot R.M.;  
RT "Primary structures of cytochrome P-450 isozyme 5 from rabbit and rat  
RT and regulation of species-dependent expression and induction in lung  
RT and liver: identification of cytochrome P-450 gene subfamily IVB.";  
RL Mol. Pharmacol. 35:617-625(1989).  
CC -1- FUNCTION: Cytochromes P450 are a group of heme-thiolate  
CC monooxygenases. In liver microsomes, this enzyme is involved in an  
CC NADPH-dependent electron transport pathway. It oxidizes a variety  
CC of structurally unrelated compounds, including steroids, fatty  
CC acids, and xenobiotics.  
CC -1- CATALYTIC ACTIVITY: RH + reduced flavoprotein + O(2) = ROH +  
CC oxidized flavoprotein + H(2)O.  
CC -1- SUBCELLULAR LOCATION: Membrane-bound. Endoplasmic reticulum.  
CC -1- INDUCTION: P450 can be induced to high levels in liver and other  
CC tissues by various foreign compounds, including drugs, pesticides,  
CC and carcinogens.  
CC -1- SIMILARITY: Belongs to the cytochrome P450 family.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; M29853; AAA41778.1; -.  
DR PIR; B40164; B40164.  
DR HSSP; P14779; 1UPZ.  
DR InterPro; IPR001128; Cytochrome\_P450.  
DR Pfam; PF00067; P450; 1.  
DR PRINTS; PR00385; P450.  
DR PROSITE; PS00086; CYCROCHROME\_P450; 1.  
DR OXidoreductase; Monooxygenase; Electron transport; Membrane; Heme;  
KW Microsome; Endoplasmic reticulum.  
FT METAL 453 453 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).  
SQ SEQUENCE 511 AA; 58936 MW; CEAFBBSBEZED944 CRC64;  
Query Match 17.6%; Score 481; DB 1; Length 511;  
Best Local Similarity 27.1%; Prod. No. 1.2e-23;  
Matches 144; Conservative 102; Mismatches 210; Indels 76; Gaps 20;  
QY 18 VILSLVTKGMSILMWRPKIEHFSKQIGRPYHFIQNYKL--VGNMLKASHMP 75

DB 20 VILMVVLKFSLL-LRQKLAR--AMSPFGPPTHMLFGALIELQGLGSDKYVSNAGQ 76  
QY 76 FSHNLIPLRVLSFYHMRKIYATLVWTFQPRR-LTVADPLLIELIFSKSEFYKNAHP 134  
DB 77 FPH-----AHLPMFGQFVGFINIYEPDYAAVYSRGD-PRAAVYD 116  
QY 135 LVKQLEGDGLSKGEKNAHHRKIISPTFHEMLKILVPLVLSKSTVDVWKSDKLSNG 194  
DB 117 PFLQWIGKLLVLDGPKFQHRKILTGFPHVDLKPYYAIAESTRWMLDMCKKABENK 176  
QY 195 EYEV-DYVWFPQILTEDVTSRTAG-----SSYEDGRAVFRIOAQOMLLCAEAF 242  
DB 177 SFDFCDVGH---VALDTMKCTGKGDGSLGHRDNGSYLAVDLTLMOQRI---DSF 229  
QY 243 Q-KVRI-----PGYRFFPTRNGLSKRLDKIRSKLKERRONAIDEGECKEPA 295  
DB 230 QYHNDFTYWLTPHGRF-----LRACKIADHTDEVI---RQKALQDEKREKKIQ 279  
QY 296 AK-DILGIMIOAKN-----VTQDIVECKSFPPAGKQTTNLTWTTLISMPWQA 348  
DB 280 RRLDPLDILLGVDBSGIKSDAELRAEVDTFWEGHDTTSGISWFLCMALYPSHQ 339  
QY 349 KARDEYLRYCGSRDVPYTDHVVKLTISMLNESRLYPVATITRAKSDVK-LGCKI 407  
DB 340 LCRSEYRGLIGDQDFOWDILAKMTYLLTMCKEFCRLYPVQYRQLNKPVPFDGSL 399  
QY 408 PGCTELLPIIAVHDOAIWQNDVNEFNPAFADGVPR-AAKHVPGFIPGLGVRTICIG 465  
DB 400 PAGSLISMTIYALHNSTYW-PDEVPFPLRNS---PENAGRHFPAMPSAGPRNCIG 455  
QY 466 QNLAILQAKTLAVMIQRTFHLAPTYQHAPVLMILYPOGAPITFRRLTN 517  
DB 456 QCFANMEMKVYALCLHREFELDPSKXPIKVPILRSKNGIHRLKPLAS 507  
RESULT 12  
YRV8 CAEEL STANDARD; PRT; 519 AA.  
ID YRV8 CAEEL STANDARD; PRT; 519 AA.  
AC Q27520;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DB Putative cytochrome P450 CYP13A1 (BC 1.14.-.-).  
GN CYP13A1 OR T10B9.8.  
OS Caenorhabditis elegans.  
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;  
OC Rhabditidae; Peloderinae; Caenorhabditis.  
OX NCBI\_TaxID=6239;  
RN [1]  
RX SEQUENCE FROM N.A.  
RA STRAIN=Briscot N2;  
RA Gardiner A.;  
RL Submitted (MAR-1995) to the EMBL/GenBank/DBJ databases.  
CC -1- FUNCTION: Cytochromes P450 are a group of heme-thiolate  
CC monooxygenases. They oxidize a variety of structurally unrelated  
CC compounds, including steroids, fatty acids, and xenobiotics.  
CC -1- SIMILARITY: Belongs to the cytochrome P450 family.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; Z48717; CAA88610.1; -.  
DR PIR; T24784; T24784.  
DR HSSP; P14779; 1UPZ.  
DR WormPep; T10B9.8; CE01660.  
DR InterPro; IPR001128; Cytochrome\_P450.  
DR Pfam; PF00067; P450; 1.

DR PRINTS, PR00385, P450.  
 DR PROSITE, PS00086; CYTOCHROME P450; 1.  
 KW Hypothetical protein; Oxidoreductase; Monooxygenase; Heme.  
 FT METAL 465 465 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).  
 SQ SEQUENCE 519 AA; 60044 MW; 3AF37EDC43539D7A CRC64;  
 Query Match 17.6%; Score 480; DB 1; Length 519;  
 Best Local Similarity 28.5%; Pred. No. 1.5e-23;  
 Matches 144; Conservative 86; Mismatches 182; Indels 94; Gaps 19;  
 QY 32 WWRPKIIEHFSKQIGRP--PYHFFIGVKELVGMMLKASHMPESHNLPRVLSFY 88  
 DB 26 FWR-----RGGVGVPMGFP-----VLGVFLNSLDNNEFPFLQC----- 58  
 QY 89 HHMKITGATFLVWFGFTPLTVADPLIREIF--SKSEPEYKNEAFLVKQLEGD--- 142  
 DB 59 REWTKKFKGKYGFTGEGTKTLVSDPLVEHVTVQYDNFYGRK-----NFIQDSKE 113  
 QY 143 ---GLSLKGEKMAHHRKIIISPTFHENMLKLVPVYLKSVTDMVKSDKSENSEVEVD 199  
 DB 114 KRTLFLPAAGFRWRKRLAISPTFSNSLAKLVQVEDSALELL--RHIEKQAGGK-QID 171  
 QY 200 VYEWFOITLEDVIRITLPGSSYEDGRAVFLQAOQML--LCAEAFQKVIIPGYPFP-- 255  
 DB 172 MLKPYQEFLLDVIRIMG--QDSQMFKNPIPIVSKLPQGNFALFLIG-GIFPYFL 227  
 QY 256 -----RGTLK--SRKLDKEIRKSLKLIERRONAIIDGEGECKEPPAKDLIGMI 304  
 DB 228 VEIRIQLIKLKLKGSFRKINETLIDAIHRIKOREED--QKNGIIGRPA--DFIDFL 283  
 QY 305 QAK-----NVTVDIYBECKSPFPAGKQTTSNLTLTWITLIS 341  
 DB 284 DAKADVEHFGENNDSKSTVTYNNRQLTTEIEIVGQTVFLIAGPDTLISLSTAYTLA 343  
 QY 342 MHPWQAKARDEVIRVCGSRDVPKDHVYLKTLMSLINESLTP--PIVATIRAKSDV 400  
 DB 344 THPEIQKLQEEVNRCEQNPV--TIDQSLKLTKECVFKALMLYLGLAFANSRRKMT 402  
 QY 401 KLGVKIPGCTELLPIIAVHDDQAIWGDVNEFNPAKADGVPRAPKAPVGFIPGLGV 460  
 DB 403 KLGMKVEVGMIDVDTWTLHTDENIGDADDEKPERMQTPNSDQIYKSGYIPGLGP 462  
 QY 461 RTICQNTALIQAKTLTAVMIGRFTF 486  
 DB 463 RQIGMRLAYMEKILLVHLIRKPTF 488  
 RESULT 13  
 CP3G\_MOUSE STANDARD; PRT; 504 AA.  
 ID CP3G\_MOUSE  
 AC Q64481;  
 DT 15-JUL-1999 (Rel. 38, Created)  
 DT 15-JUL-1999 (Rel. 38, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 GN Cytochrome P450 3A16 (BC 1.14.14.1) (CYP1A16).  
 OS CYP3A16 OR CYP3A-16  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_Taxid=10090;  
 RN (1)  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=ICR; TISSUE=Liver;  
 RX MEDLINE=95112853; PubMed=7813478;  
 RA Itoh S.; Setoh W.; Abe Y.; Hashimoto H.; Yanagimoto T.; Kametaki T.;  
 RT "A novel form of mouse cytochrome P450 3A (Cyp3a-16). Its cDNA cloning  
 and expression in fetal liver.";  
 RL Eur. J. Biochem. 226:877-882(1994).  
 CC FUNCTION: Cytochromes P450 are a group of heme-thiolate  
 CC monooxygenases. In liver microsomes, this enzyme is involved in an  
 CC NADPH-dependent electron transport pathway. It oxidizes a variety  
 CC of structurally unrelated compounds, including steroids, fatty  
 CC acids, and xenobiotics.

CC -| CATALYTIC ACTIVITY: RH + reduced flavoprotein + O(2) = ROH +  
 CC oxidized flavoprotein + H(2)O.  
 CC -| SUBCELLULAR LOCATION: Membrane-bound. Endoplasmic reticulum.  
 CC -| DEVELOPMENTAL STAGE: Fetal- and puberty-specific.  
 CC -| INDUCTION: P450 can be induced to high levels in liver and other  
 CC tissues by various foreign compounds, including drugs, pesticides,  
 CC and carcinogens.  
 CC -| SIMILARITY: Belongs to the cytochrome P450 family.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; D26137; BA05133.1; -  
 DR PIR; S50892; S50892.  
 DR HSBP; P14779; IOPZ.  
 DR MGD; MGI:106099; Cyp3a16.  
 DR InterPro; IPR001128; Cytochrome P450.  
 DR InterPro; IPR008072; BP450\_CYP3A.  
 DR Pfam; PF00067; P450; 1.  
 DR Pfam; PF00669; BP450IICYP3A.  
 DR PRINTS; PR00385; P450.  
 DR PROSITE; PS00086; CYTOCHROME P450; 1.  
 DR Oxidoreductase; Monooxygenase; Electron transport; Membrane; Heme;  
 KW Oxidoreductase; Monooxygenase; Electron transport; Membrane; Heme;  
 KW Microsome; Endoplasmic reticulum.  
 FT METAL 443 443 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).  
 SQ SEQUENCE 504 AA; 57929 MW; E364B91FD10D5D1A CRC64;  
 Query Match 17.6%; Score 479.5; DB 1; Length 504;  
 Best Local Similarity 28.2%; Pred. No. 1.5e-23;  
 Matches 145; Conservative 86; Mismatches 175; Indels 107; Gaps 18;  
 QY 18 VILSLVVKMSILMWRPKIEHFSKQIGRPYHFFIGVKELVGMMLKASHMPSPS 77  
 DB 13 VLAILIV-----LLYVGYTTHGLFKKQIGIPG-----KPLPF- 46  
 QY 78 HNLPRVLSYTH--WR-----KYGATFLVWFGFTPLTVADPLIREIFSK-----SEY 127  
 DB 47 ---LGYLVANYKGLMFWDECYEYKGTGLFEGQIPLFVITDPEITIKVLYKCSVYT 103  
 QY 128 EKNEAHPLVKQSLGGDGLS-----LKGEKMAHHRKIIISPTFHENMLKLVPVYLKSVTDM 182  
 DB 104 NRQDFPV-----GIMSKSISLADDEMKRYRLALSTPTFSNKLKEMFVIL--EQYDI 155  
 QY 183 VDKMSDLSNGEVEVDVYEWFOITLEDVIRITLPGSSYEDGRAVFLQAOQML--LCAEAFQKVIIPGYPFP-- 255  
 DB 156 LVKTLRQBAEKGK--PVAVKDVLGAYSMVDVISTTFGVNIDSINNPEDPVEVNAKRYLFD 214  
 QY 222 AQQMLLAERAFQKVFIPGRV-----FPPRGNLKSRLKDEIRKSLKLIERRONAIIDG 286  
 DB 215 YFDPLSLVALFPPLTIYMLNICMF-----KDSIEFPKTVDMTEKRLDS 263  
 QY 287 EGEECKEPPAKDLIGMIQAKN-----VTVDIYBECKSPFPAGKQTTSN 331  
 DB 264 -----KQHRVDPIYLMEAYNKSCKDSKSLSEIEITPAQSIIT-----FIFAGYETTS 313  
 QY 332 LLTWITLILSHRPMQAKARDEVIRVCGSRDVPKDHVYLKTLMSLINESLTP--PIVIA 391  
 DB 314 ILSTFVSLATHPIQKQLEIDEALPNAPPTIYTVAMETIDVNLATLALYITNR 373  
 QY 392 TIRAPKSPVKGKYGKIPGCTELLPIIAVHDDQAIWGDVNEFNPAKADGVPRAPKAPV 451  
 DB 374 LQVCKQKVEINGIYIPKSGTVIIPSVLHDDQHM--PEDEEPOPERFSK--ENKGSIDPY 431  
 QY 452 GIPFGGLGVRCICQNTALIQAKTLTAVMIGRFTF 486  
 DB 432 VLPFGNGPNCIGMRPALMMKALIKVLQNSF 466



DR Pfam: PF00067; P450; 1.  
 DR PRINTS: PR00385; P450.  
 DR PROSITE: PS00086; CYTOCHROME\_P450; 1.  
 KW Oxidoreductase; Monooxygenase; Electron transport; Membrane; Heme;  
 KW Microsome; Endoplasmic reticulum.  
 FT METAL 453 453 IRON (HEME AXIAL LIGAND).  
 FT CONFLICT 37 37 R -> Q (IN REF. 2).  
 SQ SEQUENCE 511 AA; 59019 MW; E24410DF0707A8F6 CRC64;

Query Match 17.5%; Score 477; DB 1; Length 511;

Best Local Similarity 27.8%; Pred. No. 2,2e-23; Indels 102; Gaps 22;  
 Matches 148; Conservative 92; Mismatches 190;

```

QY 10 IPKVLVLS-----VILSLVYKMSLWMPRKIEHFSKQIRGPPYHFIQNV 59
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
DB 2 VPSFLSLSPSSLSIGMASGLITVGFKLHL-LRRRTLAKAMDK--PPGPTHWLFGHA 58
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
QY 60 KEL--VGMMLKAS--HPMPSHNILPRVLSFYHHKRYGATFLVMGPTFR-LTVADP 114
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
DB 59 LEIQTGSLDKVSMHQPPYAPL-----WFGQFTGFLNTIYEP 97
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
QY 115 DLIRFISKSEFYKNAHPVVKQLEGDGLSLKGEKMAHKKIISPTFMENLKILVAV 174
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
DB 98 DYAAVYSRSD-PKAPDYDPFLQWIGRGLVLEGPKMLQHRKLLPGFHYDLKPYAV 156
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
QY 175 VLKSVTDVMDKMSDKLSENGEVEV--DYEMFQILTEDIYSTAFG-----SSTE 222
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
DB 157 FTETRIWLDCKEKKAREKSPDIFCDVGH---MALNTLMKCTFGRGDTGLGHRDSSY 212
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
QY 223 DGRAVRLQAQOMLCAEAPQKVF-----PGYRF-----PTRGNLKSRL--DK 266
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
DB 213 LAVSDLLMQQLVSPQ-YHNDFTYLTGGRFLPACQVAHDHTDQVIRERKALQDE 271
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
QY 267 EIRKSLKLTERRRQNAIDGSECKEPAKDLLGLMTQAKN---VTVDIVBECKSF 321
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
DB 272 KVRKK---IQNRH-----LDFLDLILGARDEDDIKLSDADLRREVDTF 312
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
QY 322 FPAKQTTSNLITTTTLLSMHPWQAKARDEVLRVCGSRDVPYTKDHYVKLTLSMILNE 381
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
DB 313 MEEGHDITTSISWFLYCMALYPFHQRCEVEIREIIGDQDFQWDDLGKMTYLTWCIE 372
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
QY 382 SLRLVPPIVATIRRAKSDVK-LGGYKIPCGTELLPIIAVHDAQMGNDVNEFNAPPA 440
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
DB 373 SFRILPVPQYRQLSKPYTFVDGRSLPAGSLISMHIYALHRSAYW-PDPEYFDSLRF 431
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
QY 441 DGVPRAAK-HPVGFIPFGLGVRTICIGNLAILQAKLTAVMIQRTFHLAPT 491
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
DB 432 --TENASKRHHPAFMFPFAGPRNCIGQPFAMSEMKVVTAMCLLRFPSLDPS 481
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||

```

Search completed: May 19, 2004, 10:26:20  
 Job time : 19 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd

OM protein - protein search, using sw model

Run on: May 19, 2004, 10:20:55 ; Search time 46 Seconds  
(without alignments)  
3566.728 Million cell updates/sec

```

Title:          US-09-992-901-2
Perfect score:  2730
Sequence:       1 MEESSGSNFIKVLVLSVL.....LTPHGAPITFRRLTNHD 520

```

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

```
Searched:      1017041 segs, 315518202 residues
Total number of hits satisfying chosen parameters: 1017041
```

```
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
```

```
Post-processing: Minimum Match 0%
                  Maximum Match 100%
                  Listing first 45 summaries
```

```

1:  sp. archaea: *
2:  sp. archaea: *
3:  sp. archaea: *
4:  sp. archaea: *
5:  sp. archaea: *
6:  sp. archaea: *
7:  sp. archaea: *
8:  sp. archaea: *
9:  sp. archaea: *
10: sp. archaea: *
11: sp. archaea: *
12: sp. archaea: *
13: sp. archaea: *
14: sp. archaea: *
15: sp. archaea: *
16: sp. archaea: *
17: sp. archaea: *

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	2722	99.7	520	10	Q48786	Q48786 arabidopsis
2	1673	61.3	589	10	Q84587	Q84587 oryza sativ
3	1404	51.4	537	10	Q8LIF2	Q8LIF2 oryza sativ
4	1187	42.8	528	10	Q8LL74	Q8LL74 zea mays su
5	1167.5	42.8	527	10	Q8LGM6	Q8LGM6 zea mays su
6	1164.5	42.7	523	10	Q9PDD1	Q9PDD1 oryza sativ
7	1154.5	42.3	525	10	Q9ANTU1	Q9ANTU1 oryza sativ
8	1153.5	42.2	525	10	Q9ANTU2	Q9ANTU2 oryza sativ
9	1147.5	42.0	525	10	Q9ANTU3	Q9ANTU3 oryza sativ
10	1147.5	42.0	525	10	Q9ANTU4	Q9ANTU4 oryza sativ
11	1138.5	41.7	525	10	Q9ANTU5	Q9ANTU5 oryza sativ
12	1126	41.2	515	10	Q9AX23	Q9AX23 oryza sativ
13	1124.5	41.2	515	10	Q9LUD2	Q9LUD2 arabidopsis
14	1124.5	41.2	515	10	Q8WLE1	Q8WLE1 arabidopsis
15	1119	41.0	528	10	Q8JGNA	Q8JGNA zea mays su
16	1113	40.8	524	10	Q42700	Q42700 catharanthus

AC	048786	PRELIMINARY	PRT	520 AA.
17	1112	40.7	511	10 Q9FED1
18	1111	40.7	531	10 Q9FED3
19	1110	40.7	516	10 Q43701
20	1110	40.7	528	10 Q9E9F0
21	1108	40.6	528	10 Q9FED6
22	1098	40.2	512	10 Q9LUD3
23	1089.5	39.9	512	10 Q9LUD6
24	1088.5	39.9	531	10 Q9FEP3
25	1080.5	39.6	535	10 Q9EPC7
26	1079.5	38.5	512	10 Q9LUD8
27	1079	38.5	544	10 Q9FEB1
28	1075.5	38.4	512	10 Q9LUD5
29	1067	39.1	565	10 Q9FEB6
30	1060	38.8	552	10 Q9E1C6
31	1050.5	38.5	505	10 Q9LUD0
32	1047	38.4	532	10 Q9FED6
33	1045	38.3	498	10 Q9M4X2
34	1040.5	38.1	519	10 Q9LHV0
35	1028.5	37.7	505	10 Q9F8K4
36	1024.5	37.5	512	10 Q9LUC9
37	1023	37.5	508	10 Q40411
38	1021.5	37.4	520	10 Q8L4Q4
39	1019	37.3	517	10 Q80729
40	1017.5	37.3	513	10 Q8E9C0
41	1015	37.2	514	10 Q94EY6
42	1006	36.8	534	10 Q9ASR3
43	999	36.6	519	10 Q9L172
44	997	36.5	430	10 Q9L172
45	990.5	36.3	518	10 Q9F8C3

ALIGNMENTS

RESULT 1

048786

ID 048786 PRELIMINARY PRT 520 AA.

AC 048786

DT 01-JUN-1998 (TREMBLrel. 06, Created)

DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)

DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)

DE Putative cytochrome P450.

GN ATG326710.

OS Arabidopsis thaliana (Mouse-ear cress).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;

OC eustroids II; Brassicales; Brassicaceae; Arabidopsis.

OX NCBI Taxid=3702;

[1]

RN SEQUENCE FROM N.A.

RP STRAIN=cv. Columbia;

RC MEDLINE=20083487, PubMed=10617197;

EX Lin X., Kaul S., Rounsley S.D., Shea T.P., Benito M.-I., Town C.D., Firth C.Y., Mason T.M., Bowman C.U., Barnstead M.E., Feldblyum T.V., Buell C.R., Kechum K.A., Lee J.D., Rongting C.M., Koo H., Moffat K.S., Cronin L.A., Shen M., Vanaken S.B., Umayam L., Tallon L.J., Gill J.B., Adams M.D., Carrera A.J., Creasy T.H., Goodman H.M., Somerville C.R., Copenhagen G.P., Preuss D., Niernm M.C., White O., Eisen J.A., Salzberg S.L., Fraser C.M., Venter J.C.;

RA "Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana."

RT thaliana."

RL Nature 402:761-768(1999).

[2]

RC SEQUENCE FROM N.A.

RP STRAIN=cv. Columbia;

RL Lin X.;

RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.

CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.

DR EMBL: AC003105; AAB95305.1; --

DR PIR: H84663; H84663.

DR HSSP: P14779; 1BVY.

DR GO: GO:0004497; F:monooxygenase activity; IEA.

DR GO: GO:0006118; P:electron transport; IEA.

## ALIGNMENTS



DR InterPro: IPR001128; Cytochrome\_P450.  
 DR Pfam: PF00067; P450; 1.  
 DR PRINTS: PR00385; P450.  
 DR PROSITE: PS00086; CYTOCHROME\_P450; 1.  
 KN Heme; Monooxygenase; Oxidoreductase.  
 SQ SEQUENCE 520 AA; 59490 MW; 585186A0CC07A59 CRC64;

Query Match 99.7%; Score 2722; DB 10; Length 520;  
 Best Local Similarity 99.8%; Pred. No. 1,8e-216;  
 Matches 519; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MESESSSWPFPKLVSVLSVLYVKGMSLLMPPRKLSEHSGKGIKGPYHFIQV 60  
 DB 1 MESESSSWPFPKLVSVLSVLSVLYVKGMSLLMPPRKLSEHSGKGIKGPYHFIQV 60  
 QY 61 ELVGMMLKASHHMPFSHNLPRVLSFYHMKRIYGAFTLVWFGPTLVADPDIREI 120  
 DB 61 ELVGMMLKASHHMPFSHNLPRVLSFYHMKRIYGAFTLVWFGPTLVADPDIREI 120  
 QY 121 FSKSEFEKNEAHPVYKQLEGDGLSLKGEKNAHRKLIISPTFHENKILVAVLSVT 180  
 DB 121 FSKSEFEKNEAHPVYKQLEGDGLSLKGEKNAHRKLIISPTFHENKILVAVLSVT 180  
 QY 181 DWVDKMSDKLSENGEVDVYEWFOITLTDVTSRTAFSGSYEDGAVFRLQAQMLCAE 240  
 DB 181 DWVDKMSDKLSENGEVDVYEWFOITLTDVTSRTAFSGSYEDGAVFRLQAQMLCAE 240  
 QY 241 AFQKVPFPGYRFPPTNGNLKSKRLDKYIKSLIKLIERROVALDGESECEPAKDL 300  
 DB 241 AFQKVPFPGYRFPPTNGNLKSKRLDKYIKSLIKLIERROVALDGESECEPAKDL 300  
 QY 301 GLMIQAKNTVODIVECKSFPAFGKQTSTNLTWTTLISMPHMOAKADEVLRVCGS 360  
 DB 301 GLMIQAKNTVODIVECKSFPAFGKQTSTNLTWTTLISMPHMOAKADEVLRVCGS 360  
 QY 361 RDVPTKHVVVKKLTIISLINESLRILYPIVATIRAKSDVKGKIPGTELLPIIAV 420  
 DB 361 RDVPTKHVVVKKLTIISLINESLRILYPIVATIRAKSDVKGKIPGTELLPIIAV 420  
 QY 421 HHQAVIWNVDNENFNPAPFADGVPRAPKIPVGPFGVGTICIGNALIIQAKTLAVM 480  
 DB 421 HHQAVIWNVDNENFNPAPFADGVPRAPKIPVGPFGVGTICIGNALIIQAKTLAVM 480  
 QY 481 IORTFHIAPTYOHAPTVMLLYPOHGAFTFRRLTNHD 520  
 DB 481 IORTFHIAPTYOHAPTVMLLYPOHGAFTFRRLTNHD 520

## RESULT 2

Q84SB7 PRELIMINARY; PRT; 589 AA.

AC Q84SB7; TREMBLrel. 24, Created)  
 DT 01-JUN-2003 (TREMBLrel. 24, Last sequence update)  
 DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)  
 DE OSUNBD0008D07.14 protein.  
 GN OSUNBD0008D07.14.  
 OS Oryza sativa (japonica cultivar-group).  
 CC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;  
 CC Euphorbia; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 CC Euphorbiaceae; Oryzae; Oryza.  
 NCBI\_TaxID=39947;  
 RN NCB1  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. Nipponbare;  
 RX MEDLINE=22337376; PubMed=12447438;  
 RA Sasaki T., Matsumoto T., Yamamoto K., Sakata K., Baba T., Katayose Y.,  
 RA Wu J., Niimura Y., Cheng Z., Nagamura Y., Antonio B.A., Kanamori H.,  
 RA Hosokawa S., Maekawa M., Arikawa K., Chiden Y., Hayashi M.,  
 RA Okamoto Y., Ando T., Aoki H., Arita K., Hamada M., Harada C.,  
 RA Hijishita S., Honda W., Ichikawa Y., Ikonuma A., Iijima M., Ikeda M.,  
 RA Ikono M., Itoh Y., Itoh Y., Itoh Y., Iwabuchi A., Kamiya K.,  
 RA Karasawa W., Katagiri S., Kikuta A., Kobayashi N., Kono I.,

RA Machita K., Maehara T., Mizuno H., Mizubayashi T., Mukai Y.,  
 RA Nagasaki H., Nakashima M., Nakama Y., Nakamichi Y., Nakamura M.,  
 RA Namiki N., Negishi M., Ohta I., Ono N., Saji S., Sakai K., Shibata M.,  
 RA Shimokawa T., Shomura A., Song J., Takasaki Y., Terasawa K., Tsuji K.,  
 RA Waki K., Yamagata H., Yamane H., Yoshiki S., Yoshihara R., Yukawa K.,  
 RA Zhong H., Iwama H., Endo T., Ito H., Hahn J.H., Kim H.I., Eun M.Y.,  
 RA Yano M., Jiang J., Gojobori T.,  
 RT "The genome sequence and structure of rice chromosome 1."  
 RL Nature 420:312-316(2002).  
 RN [2]

RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. Nipponbare;  
 RA Sasaki T., Matsumoto T., Katayose Y.,  
 RL Submitted (FEB-2003) to the EMBL/Genbank/DBJ databases.  
 DR EMBL:AB006237; BAC57879.1; -  
 DR GO:0006118; P:electron transport; IBA.  
 DR InterPro: IPR001128; Cytochrome\_P450.  
 DR Pfam: PF00067; P450; 1.  
 DR PRINTS: PR00385; P450.  
 DR PROSITE: PS00086; CYTOCHROME\_P450; 1.  
 SQ SEQUENCE 589 AA; 66070 MW; 23892D1FB457DFA8 CRC64;

Query Match 61.3%; Score 1673; DB 10; Length 589;  
 Best Local Similarity 56.6%; Pred. No. 1.3e-129;  
 Matches 318; Conservative 87; Mismatches 79; Indels 78; Gaps 8;

QY 23 VIVKMSLLMPPRKLSEHSGKGIKGPYHFIQVYVGMMLKASHHMP--FSHN 79  
 DB 17 VAVKLEVMMPRVEEHFAQGITGRYRFLVGVREVALVVAASAKMPPYRSHN 76  
 QY 80 ILPRVLSFYHMKRIY----- 95  
 DB 77 VLPRVLAIFYHMKRIYGVNPPPLNLNLSIQOKQPRTRMQVAVGERRAPGYDIDM 136  
 QY 96 ---GATFIVMPGPTFRLVADPDIREI-FSKSEFEKNEAHPVYKQLEGDGLSLNGEK 151  
 DB 137 AALGTFPLWFGPTRLADBELIREVILARADFDESHPMVRQEGSLVSRGDK 196  
 QY 152 WAHHRKIISPTFHENKILVAVLSVTDVDK---SDKSENGEVDVYEWFOI 206  
 DB 197 WAHHRVLPAPFHMNLRLILPCVGMTVLDMAKRRANABDK--SEVEIDVSDWFOV 253  
 QY 207 LTFEDVTSRTAFSGSYEDGAVFRLQAQMLCAEAFQVFIPIGYRFPPTNGNLKSKRLDK 266  
 DB 254 VIEDALTRAFSGSYEDGAVFRLQAQMLCAEAFQVFIPIGYRFPPTNGNLKSKRLDK 313  
 QY 267 EIKSLILKLIERRONALDGESECEPAKDLGLMTQA-----KNTVQ 312  
 DB 314 EIKNLVTLIGRQEGDDEKLDG---AKDLGLMTNAAASNGRSLPVPSPITVN 369  
 QY 313 DIVECKSFPAFGKQTSTNLTWTTLISMPHMOAKADEVLRVCGSDVPTKHVVYL 372  
 DB 370 DIVECKTFPAFGKQTSTNLTWTTLISMPHMOAKADEVLRVCGSDVPTKHVVYL 429  
 QY 373 KTLISLINESLRILYPIVATIRAKSDVKGKIPGTELLPIIAVHDDQIWNVDV 431  
 DB 430 KTLISLINESLRILYPIVATIRAKSDVKGKIPGTELLPIIAVHDDQIWNVDV 489  
 QY 432 NENPAPFADGVPRAPKIPVGPFGVGTICIGNALIIQAKTLAVMIRFTHLAPT 491  
 DB 490 AQPENPAPFADGVPRAPKIPVGPFGVGTICIGNALIIQAKTLAVMIRFTHLAPT 549  
 QY 492 YOHAPTVMLLYPOHGAFTFR 513  
 DB 550 YVHAPTVMLLYPOHGAFTFR 571

## RESULT 3

Q84IF2 PRELIMINARY; PRT; 537 AA.

AC Q84IF2; TREMBLrel. 22, Created)  
 DT 01-OCT-2002 (TREMBLrel. 22, Last sequence update)

DT 01-OCT-2003 (Tremblrel. 25, last annotation update)  
 DE Putative cytochrome P450.  
 GN OJ1316\_A04.10 OR P0503D09.27.  
 OS *Oryza sativa* (japonica cultivar-group).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 OC Ehrharioideae; Oryzaceae; *Oryza*.  
 NCBI\_Taxid=39947;  
 RN (1)  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. Nipponbare;  
 RA Sasaki T., Matsumoto T., Yamamoto K.;  
 RT *Oryza sativa* nippobare(GM3) genomic DNA, chromosome 7, BAC  
 RT clone:OJ1316\_A04.10;  
 RN Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. Nipponbare;  
 RA Sasaki T., Matsumoto T., Katayose Y.;  
 RT *Oryza sativa* nippobare(GM3) genomic DNA, chromosome 7, PAC  
 RT clone:P0503D09.27;  
 RN Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.  
 CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.  
 DR EMBL, AP003822; BAC06993.1; -.  
 DR EMBL, AP005455; BAC16737.1; -.  
 DR Gramene; OBLIR2; -.  
 DR GO, GO:0004553; F:hydrolyase activity, hydrolyzing O-glycosyl . . .; IEA.  
 DR GO, GO:0004497; F:monooxygenase activity; IEA.  
 DR GO, GO:0005975; P:carbohydrate metabolism; IEA.  
 DR GO, GO:0006118; P:electron transport; IEA.  
 DR InterPro; IPR001128; Cytochrome P450.  
 DR InterPro; IPR008263; Glyco\_hyd16\_AS.  
 DR Pfam; PF00067; P450; 1.  
 DR PRINTS; PR00385; P450.  
 DR PROSITE; PS00086; CYTOCHROME\_P450; 1.  
 DR PROSITE; PS01034; GLYCOSYL\_HYDROL\_F16; 1.  
 DR Heme; Monooxygenase; Oxidoreductase.  
 SQ SEQUENCE 537 AA; 59909 MW; 068441D7D046ED20 CRC64;

## Query Match

Best Local Similarity 51.4%; Score 1404; DB 10; Length 537;  
 Matches 271; Conservative 97; Mismatches 138; Indels 24; Gaps 10;

QY 7 SMFIPKVLVLSVLSLVV-----KMSLWMPRKIEHFSKQIGRPYHFFIGNVKE 61  
 DB 10 SMSGAAVAVAAMAAVVAARAAALWMPRRVRRHRAAGVGPGVRFVSGSIS 69  
 QY 62 LVGMMLKASSHMP--PESHNIPRLSTFYHMKIKYCATLVWGFRTLVADPDILIRE 119  
 DB 70 LVRLVNDASRMEPPTSHDILPRVLPYHMRKLYGPMILWGRTPRLVSPFELIRE 129  
 QY 120 I-FSKSEFEKNEAPLVKOLEGDLISLKGKNAHRIKISPTFHEHNLKLVAVYVLS 178  
 DB 130 VLLTRADHDFREARPMICQPEGYLSLHGERNARRRRLTPAFHENTRMIAFVAGT 189  
 QY 179 VTDVMDKMSDKLSEN--GEVEVDYEMQIITLSDVISTARG--SSYEDGAIVRLAQOM 225  
 DB 190 VTRMDDELAEARAGAGAEVDVEMFORVDEAITFAAGRRNYDDGAIVRLQDELTA 249  
 QY 236 LICAERFQVFIPIGRFFPTTRGNLSRKLDKEIRKSLKILERRONAIGEGSECKE-- 293  
 DB 250 GYATAAHSKVITIPGRFLPTRKNRWLDREINSHLAKF--TGLQSSSHDDDDG 308  
 QY 294 ---PAKDLGLMIOAKNVVDIVIECKSPFPAGKQTTSMILTWITILLSPENQAKA 350  
 DB 309 DGGGGRBEMSMFAA--MTAGEIIEESKNEFFAKETLSNLTWITVALAMPWQERA 366  
 QY 351 RDEVLRYGSSRDVPTKDVVVKLTSMILNESLRLPYIVATIRASDVYLGQYKIPCQ 410  
 DB 367 KREVAVACGRDPLPKDHLPKLKITGLINLETSLRYPYAVAMITAEDEVLEGGCVAPAG 426  
 QY 411 TELLPITIAVHHDAQIWNQDVNEFNPAFA--DGVPAAGVPGVIFPGIGVYTCIGON 467

DB 427 TEWPIPIAVVHHDAAMGDDAAEFNPARFAADDGGR--RHPMFMEGGGARVCIQON 484  
 QY 468 LAIQAKITLAWMIQRTFFHLAPYQAPVLMILYPOGAPITTRRLTN 517  
 DB 485 MALMEAKVALAVLRRREFRISPAVYAPRVLMLTSPQGAPIFRLLTS 534

## RESULT 4

QBL174  
 ID QBL174 PRELIMINARY; PRT; 528 AA.

DT 01-OCT-2002 (Tremblrel. 22, Created)  
 DT 01-OCT-2002 (Tremblrel. 22, last sequence update)  
 DT 01-JUN-2003 (Tremblrel. 24, last annotation update)  
 DE Cytochrome P450 monooxygenase CYP72A16.  
 OS Zea mays subsp. mays (maize).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 OC PACAD clade; Panicoidae; Andropogoneae; Zea.  
 NCBI\_Taxid=4578;  
 RN (1)  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. Indred 873;  
 RA Wang J., Schuler M.A.;  
 RT "Molecular Characterization of the Maize CYP71C3 and CYP72A subfamily  
 RT genes";  
 RN Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.  
 CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.  
 DR EMBL, AF455265; AA077716.1; -.  
 DR GO, GO:0004497; F:monooxygenase activity; IEA.  
 DR GO, GO:0006118; P:electron transport; IEA.  
 DR InterPro; IPR001128; Cytochrome P450.  
 DR Pfam; PF00067; P450; 1.  
 DR PRINTS; PR00385; P450.  
 DR PROSITE; PS00086; CYTOCHROME\_P450; 1.  
 DR Heme; Monooxygenase; Oxidoreductase.  
 SQ SEQUENCE 528 AA; 60642 MW; 95046F1C7112CCD CRC64;

## Query Match

Best Local Similarity 43.5%; Score 1187; DB 10; Length 528;  
 Matches 229; Conservative 99; Mismatches 177; Indels 18; Gaps 4;

QY 4 ESSSFIPKVLVLSVLSLVVKMSLWMPRKIEHFSKQIGRPYHFFIGNVKELV 63  
 DB 13 EASPARAGATAAVLVVWLAAMTLEWAMWTPRLDRLAAGLKGRTYLLTGDVRENA 72  
 QY 64 GMLLKASSHMPF--SHNIPRLSTFYHMKIKYCATLVWGFRTLVADPDILIRE 122  
 DB 73 RLNRBAKTKPLPLSHDILPRVLPFNAYKENGNTSFTWFGRLPYIIPDLNREVLVS 132  
 QY 123 KSEFEKNEAPLVKOLEGDLISLKGKNAHRIKISPTFHEHNLKLVAVYVLSVTM 182  
 DB 133 NKFGFGKPLPSRSGKLANLANBEBKAKHRIIMPFAHNEKIKGMLPVATGCCADM 192  
 QY 183 VDKMSDKLSENGEVEVDVYEFQILTLDVISTRTAFSSSYEDGAIVRLAQOMLCAEAF 242  
 DB 193 INRWKNSNSKPEBMDVWPFQVLTGDVISTRTAFSGNYDEGRNIFQLQEOAERLIQSF 252  
 QY 243 QKVPFPGYRFPPTGNLSRKLDKEIRKSLKILERRONAIGEGSECKEPAKDLGL 302  
 DB 253 QTIIPGYWFPPTKNNRMEKIDREIRKLHGILIRKKEAFIDSEGN-----DDLGL 306  
 QY 303 MIOAK-----NTVQDIVIECKSPFPAGKQTTSMILTWITILLSNHPEWQAKARD 352  
 DB 307 LVESNMRSSNNAKLGWTTEDILIECKLTFYAGAGETTSVLLTWITILLSNHPEWQAKAR 366

QY 353 EYLRCGSRDVPFKDHYVVKLTSLMINSRLRYPPIVATIRRAKSDVKLGKYPICGTE 412  
 DB 367 EVLNHNG-MGTPDPDLNRLKTYMTLYLALYPPVFLSKRTYKEMELGKIPSSVS 425  
 QY 413 LILPIIAVHHQAIWGNVNEPNAFADGVPRAAKHVGFIPGLGVRTICIGONLALIQ 472  
 DB 426 LILPIIFIHDPINWCKDASEFNPQRFEDGINSATYHQAFFPFGWGPRIICIGONFALLE 485  
 QY 473 AKTLTAVMTQRFTHLAPTYOHAPTVLMLYPOHGAFTIRRLT 515  
 DB 486 AKKALSTILQRFSELSSTYHAPYVITLHPHQAQIRLKL 528

## RESULT 5

Q8LGM8 PRELIMINARY; PRT; 527 AA.  
 AC Q8LGM8; 01-OCT-2002 (TREMBlrel. 22, Created)  
 DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)  
 DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)  
 DE Cytochrome P450 monooxygenase CYP72A5.  
 OS Zea mays subsp. mays (maize).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 OC PACCAD clade; Panicoideae; Andropogoneae; Zea.  
 NCBI\_TaxID=4578;  
 [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. Inbred B73;  
 RA Wang J., Schuler M.A.;  
 RT "Molecular characterization of the maize CYP71C3 and CYP72A subfamily  
 genes.";  
 RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.  
 [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. Inbred B73;  
 RA Wang J., Schuler M.A.;  
 RT "Xenobiotics regulate maize cytochrome P450 monooxygenases at a post-  
 transcriptional level.";  
 RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.  
 CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.  
 DR EMBL; AY072300; AAL66770.1; -;  
 DR GO; GO:0004497; F:monooxygenase activity; IEA.  
 DR GO; GO:0006118; P:electron transport; IEA.  
 DR InterPro; IPR001128; Cytochrome\_P450.  
 DR Pfam; PF00067; P450; 1.  
 DR PRINTS; PR00385; P450.  
 DR PROSITE; PS00086; CYTOCHROME\_P450; 1.  
 KW Heme; Monooxygenase; Oxidoreductase.  
 SQ SEQUENCE 527 AA; 60461 MW; 2271A93503f1P912 CRC64;

Query Match 42.8%; Score 1167.5; DB 10; Length 527;  
 Best Local Similarity 42.9%; Pred. No. 7.2e-88;  
 Matches 226; Conservative 102; Mismatches 180; Indels 19; Gaps 5;  
 QY 1 MEESSSMFIPKVLVSVLSTLVYVKGMSLLMWRPKRIEHSKQGINRPPHYFIGNVK 60  
 DB 8 MLREVPSPALASV-VASVSLMLVMTLEWAMTTPRLERARLVQGLKSTRRLFTGLDR 66  
 QY 61 ELVGMATKASGHPMP-SNITLPRVLSFYHNRKIYGATFLVFGFTFLVADDLIRE 119  
 DB 67 ETARANREARKKPLPLGSHDIARVOPMHSITKEYGKSLFTWFGFTPRVMTPDDELVE 126  
 QY 120 IFSKSEFYKNEAHPLVKOLEGGLSLKGEKNAHRIKIIPTFHENIKLLVPLVLSV 179  
 DB 127 VLSNKHFGKPRNRIGRLANGLVHDEKKAHRIINPAFHHEKIKGMPPVSTCC 186  
 QY 180 TDVYKMSDKLSENGEVEVDYEMFOLLTEDVSRFAFGSISDGAAYRLDAQOMLCA 239  
 DB 187 IEMITRMDNSMPGSSSEIDVWPFQVLTDVLSKRAFGSNYOGRRIFELQELAERLI 246  
 QY 240 EAFKVFIPGYRFFPTGNLKSRLDKRIKSLKLTIERROVAIDGEGECKEPAADL 299

DB 247 QGVQITIFIPGYWLEFPTKNNRRRAIDVEIRKLEIIIGREKDT-----KNRETNDDL 300  
 QY 300 LGLMIOAK-----NTVODIVECKSPFAGKOTSNLTMTTILSMPEWQAK 349  
 DB 301 LGLLESNTROGNSAGSLTTEDEVIECKLFYRAGHETTSVLTWTLIVLSNHPWQER 360  
 QY 350 ARDEVLRVCGSRDVPFKDHYVVKLTSLMINSRLRYPPIVATIRRAKSDVKLGKYPIC 409  
 DB 361 ABEVLSHNG-RTTDPYDSLRLKLTITMILEVLRLYPPATFLRRYKEMELGKIXPA 419  
 QY 410 GTRELPIIAVHHQAIWGNVNEPNAFADGVPRAAKHVGFIPGLGVRTICIGONLA 469  
 DB 420 GVDLLEPFIHDPINWCKDASEFNPQRFEDGINSATYHQAFFPFGWGPRIICIGONF 479  
 QY 470 ILQAKTLTAVMTQRFTHLAPTYOHAPTVLMLYPOHGAFTIRRLT 516  
 DB 480 LLEAKWTICTILQRFSELSSTYHAPYVITLHPHQAQIRLKL 526

## RESULT 6

Q9FDZ1 PRELIMINARY; PRT; 523 AA.  
 AC Q9FDZ1; 01-MAR-2001 (TREMBlrel. 16, Created)  
 DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)  
 DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)  
 DE Putative cytochrome P450.  
 GN P0688A04.10 OR P0006C01.25 OR RICE2.  
 OS Oryza sativa (Rice), and  
 OS Oryza sativa (japonica cultivar-group).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 OC Ehrhartoideae; Oryzoae; Oryza.  
 NCBI\_TaxID=4530; 35947;  
 [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. Nippondare;  
 RA Sasaki T., Matsumoto T., Yamamoto K.;  
 RT "Oryza sativa nippondare (GA3) genomic DNA, chromosome 1, PAC  
 clone: P0688A04.";  
 RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.  
 [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. Nippondare;  
 RA Sasaki T., Matsumoto T., Yamamoto K.;  
 RT "Oryza sativa nippondare (GA3) genomic DNA, chromosome 1, PAC  
 clone: P0006C01.";  
 RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.  
 [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. Nippondare;  
 RA Imahishi H., Shimizu M., Ohkawa H.;  
 RT "Cytochrome P450 in rice.";  
 RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.  
 CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.  
 DR EMBL; AF002839; BAB19104.1; -;  
 DR EMBL; AF002744; BAB19083.1; -;  
 DR EMBL; AB047400; BAB85117.1; -;  
 DR Gramene; G9FDZ1; -;  
 DR GO; GO:0004497; F:monooxygenase activity; IEA.  
 DR GO; GO:0006118; P:electron transport; IEA.  
 DR InterPro; IPR001128; Cytochrome\_P450.  
 DR Pfam; PF00067; P450; 1.  
 DR PRINTS; PR00385; P450.  
 DR PROSITE; PS00086; CYTOCHROME\_P450; 1.  
 KW Heme; Monooxygenase; Oxidoreductase.  
 SQ SEQUENCE 523 AA; 59873 MW; E31AF21BC17A83CA CRC64;

Query Match 42.7%; Score 1165.5; DB 10; Length 523;  
 Best Local Similarity 44.1%; Pred. No. 1e-87;  
 Matches 231; Conservative 97; Mismatches 175; Indels 21; Gaps 7;  
 QY 5 SSSWTFIPKVLVSVLSTLVYVKGMSLLMWRPKRIEHSKQGINRPPHYFIGNVKELV 63

```

DB      8 ASQWILAAAAAVALMLAVSTLEMAWMTPRRLERALLRAQIGRGNRYLFTGDVENV 67
QY      64 GMLLKASHMPFSS-HNILEPVLSFYHWRKRYGATFLWFGPTFRLTVADPDIREIFS 122
DB      68 RLNRBRKRPFLGCHDITPRVLPMSKAVBEHGRPSFTWGPTRWMSISPEIRVMS 127
QY      123 -KSEFEKNEAHPVYQLEGDGLSLKGRMAHRRKISPTFMENLKLVVYVLSKVD 181
DB      128 NFEHGYKRPRLGKLL-ASGVSYEGEKMAKHRIINPAFHEKIKEMLPVSNCTE 186
QY      182 MYDKMSDKLSENGEVEVDVYEWFOITTEVTSRTAFGSSYEDGRAVFRLOAQMLLQAA 241
DB      187 MYTRWNSWISBMSSEVDVWPEPQNLTDGVLSKIFGSSYEGRRIFQLOMSAERLQA 246
QY      242 FQKVPIDGYRFFPTRGNLKSRKLDKIRKSLKLTERRQNAIDGEGECKEPAKDLG 301
DB      247 FRTIFPGWFLPTKNNRRLREIREVSLLKLGIGKEBAIKOG-----ETNGDLG 300
QY      302 LMIQAK-----NVTVDIVECKSFPPAGKQTSNLTWTTTILSHMPWOKAR 351
DB      301 LIVESNMBESNGKALGMTTDEIBECKLFYAGMETTSVLLTWLTVLSHMPWQERAR 360
QY      352 DEVLRYCGSDVPTKDHVVKLKTLSMILNESLTPPIVATIRAKSDVKLGKYPGCT 411
DB      361 EYLWHRFG-RITPDVDSLSRKIVMILYEVRLVLPVPLTRTYKEMELGKYPABV 419
QY      412 ELIPIIAVHDQAIWGNVDNENPAPADGVPRAKAPVGFIPGLGVTCTICQNTAIL 471
DB      420 TMLPILFIHDPIDIGKAGEFNPGREFADGISNAKTQTSFFPGMGPRICIQNALL 479
QY      472 QAKLTAWMIQRPFTPLAPTYQAPVLMVLYPOGAPITPRL 515
DB      480 EAKMAICTTIQRPSELSPTHTAPFTVITLHPQAGQIKAKKI 523

```

```

RESULT 7
Q9ATU3 PRELIMINARY; PRT; 525 AA.
AC Q9ATU3;
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
DE Putative cytochrome P450.
OS Lolium rigidum.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;
OC Poaceae; Lolium.
OX NCBI_TaxID=89674;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SLR 31;
RA Fischer T.C., Klattig J.T., Gierl A.;
RT "A general cloning strategy of divergent plant cytochrome P450 genes
and its application in Lolium rigidum and Oculum basilicum.";
RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
CC -I- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
DR EMBL: AF321868; AAK38092.1; -.
DR GO: GO:0004497; F:monooxygenase activity; IEA.
DR GO: GO:0006118; P:electron transport; IEA.
DR InterPro: IPR001128; Cytochrome_P450.
DR Pfam: PF00067; P450; 1.
DR PRINTS: PR00385; P450.
DR PROSITE: PS00086; CYTOCHROME_P450; 1.
KW Heme; Monooxygenase; Oxidoreductase.
SQ SEQUENCE 525 AA; 59998 MW; DDFC8B96DFFP060C CRC64;

```

```

Query Match 42.3%; Score 1154.5; DB 10; Length 525;
Best Local Similarity 43.9%; Pred. No. 8.5e-87;
Matches 232; Conservative 97; Mismatches 175; Indels 25; Gaps 9;

```

```

DB      8 MIREASPNLSLGAANAALLWLAAYE---WAMTRRLRLALQAGLACTQYRLFTGD 64
QY      59 VKELVGMMLKASHMPFSS-HNILEPVLSFYHWRKRYGATFLWFGPTFRLTVADPDLI 117
DB      65 VERNRLNRBRKRPFLGCHDITPRVLPMSKAVBEHGRPSFTWGPTRWMSISPEIRVMS 124
QY      118 REIFS-KSEFEKNEAHPVYQLEGDGLSLKGRMAHRRKISPTFMENLKLVVYVVL 176
DB      125 REVLNKKGHYKQKQSSRLGTL-ANGIANHGRKMAKHRIINPAFHEKIKEMLPVPA 183
QY      177 KSVTMDVMSDKLSENGEVEVDVYEWFOITTEVTSRTAFGSSYEDGRAVFRLOAQML 236
DB      184 TCCEMITRWNDSMTBESSBIDIMPEQNLTDGVISRTAFSGVYQGMKIFQLEGELAE 243
QY      237 LCAEAPQKVPIDGYRFFPTRGNLKSRKLDKIRKSLKLTERRQNAIDGEGECKEPA 296
DB      244 RLIMAFOITIFPGWFLPTKNNKMRALDCEIR-TIIRVITRKQKAIKN-----GAIS 297
QY      297 KDLGLMIQAK-----NVTVDIVECKSFPPAGKQTSNLTWTTTILSHMPW 346
DB      298 DDLGLLESNMBESNGKADLGWSTEBIIOECKLFYAGMETTSVLLTWLTVLSHMPW 357
QY      347 QAKADEVLRYCGSDVPTKDHVVKLKTLSMILNESLTPPIVATIRAKSDVKLGKYG 406
DB      358 QEKADEVLWHRFG-RITPDVDSLSRKIVMILYEVRLVLPVPLTRTYKAMELGKIK 416
QY      407 ICGTELLIPIIAVHDQAIWGNVDNENPAPADGVPRAKAPVGFIPGLGVTCTICQ 466
DB      417 YPAGVNLMPILFIHDPIDIGKAGEFNPGREFADGISNAKHGSPFGGGRICIG 476
QY      467 NLAIIQAKLTAWMIQRPFTPLAPTYQAPVLMVLYPOGAPITPRL 515
DB      477 NFALIEAKMALSTIIQHPSELSPTHTAPFTVITLHPQAGQIKAKKI 525

```

```

RESULT 8
Q9ATU2 PRELIMINARY; PRT; 525 AA.
AC Q9ATU2;
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
DE Putative cytochrome P450.
OS Lolium rigidum.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;
OC Poaceae; Lolium.
OX NCBI_TaxID=89674;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SLR 31;
RA Fischer T.C., Klattig J.T., Gierl A.;
RT "A general cloning strategy of divergent plant cytochrome P450 genes
and its application in Lolium rigidum and Oculum basilicum.";
RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
CC -I- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
DR EMBL: AF321869; AAK38093.1; -.
DR GO: GO:0004497; F:monooxygenase activity; IEA.
DR GO: GO:0006118; P:electron transport; IEA.
DR InterPro: IPR001128; Cytochrome_P450.
DR Pfam: PF00067; P450; 1.
DR PRINTS: PR00385; P450.
DR PROSITE: PS00086; CYTOCHROME_P450; 1.
KW Heme; Monooxygenase; Oxidoreductase.
SQ SEQUENCE 525 AA; 59631 MW; ZC3032188914DF49 CRC64;

```

```

Query Match 42.2%; Score 1151.5; DB 10; Length 525;
Best Local Similarity 43.7%; Pred. No. 1.5e-86;
Matches 229; Conservative 96; Mismatches 178; Indels 21; Gaps 7;

```

```

QY      4 EGSNFIKPVTVLSVILSVIVKMSLMMRPKRIEHSQGRGPPHYHFIQNVYELV 63
DB      11 EASPNMLCAAAAVAVMLA-AWILEWAMWPRRLGRALLEQKGRIRLFTGDVENV 69

```

QY 64 GMLTKASSHPMPF-SHNILPRVLSFYHHMKRYGATFLVWGPPFRLTVADPDILREIFS 122  
 Db 70 RANKEARSKPLPLGSHDIIIPVQPMISNAIKENGLSTFWGPEPRVTIILDPESVREILS 129  
 QY 123 -KSEFEYKNEAPLVKQLEGGDILSLKGEKMAHRRKIISPFFHEMLKLVLPVYLKSVTD 181  
 Db 130 NKGHGXPRSSRFGLL-ANGLVNHQGEKMAKRRILNPAPFHEKIKRMLPVFSACSEE 188  
 QY 182 MYDKMSDKLSENGEVEVDVYEWQILTEDVTSRTAFSGSYEDGAVFRLQAQMLLCAEA 241  
 Db 189 MITRWNMSQGVSEVDVWPEFQULTGDIVISRTAFSGSYOEGTKIFOLQOGBRLMOA 248  
 QY 242 FQKVPIPGVFPFPTGNLKSRLDKERKSLKLIERRONALIDEGECKEPAKDLG 301  
 Db 249 FQTLFPGVFLPTKNRRMRREIDREICTILRGILIEKORAIKSGEA-----SSDDLIG 302  
 QY 302 LMIQAK-----NTVODIVECKSPFFAGKOTSNLLTWTIILSNHPEWQAKAR 351  
 Db 303 LLBSNRRESNGKADLGSTEDLIECKLFYFAGMETTSVLLTWTLLVLSHWPWQEQAR 362  
 QY 352 DEVLAVCGSDVPTKOHVVKLITSLMINESLRYPPIVATIRAKSDVKLGKXIPCGT 411  
 Db 363 KEVLHFFG-RTPDPENLSRLKIYVWLYEVLRLPRAIFVTRRYKAMELGITTPAGV 421  
 QY 412 ELIPIAVHDOAIMGNDVNEFPARFADGVPRAKHPVGFIPFGLGVRTICIGNTAIL 471  
 Db 422 NMLPILFIHDPNWKDASEFPORFADGISNAVXHPAAFPFGGPRICIGONFALL 481  
 QY 472 QAKTLAVMIQRTFLAPTYQAPVLMILYPOGATITRRRL 515  
 Db 482 EAKMALSTILQRFSELSPSYTHAPYVTLHPQGAIVLRKI 525

RESULT 9

Q9ATU1 PRELIMINARY; PRT; 525 AA.  
 ID Q9ATU1  
 AC Q9ATU1  
 DT 01-JUN-2001 (TREMBlrel. 17, Created)  
 DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)  
 DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)  
 DE Putative cytochrome P450.  
 OS Lolium rigidum.  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae;  
 OC Poaceae; Lolium.  
 OC NCBI\_TaxID=89674;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=SR\_31;  
 RA Fischer T.C.; Klattig J.T.; Gierl A.;  
 RT "A general cloning strategy of divergent plant cytochrome P450 genes  
 and its application in Lolium rigidum and Octium basilicum";  
 RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.  
 CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.  
 DR EMBL; AF321870; AAK38094.1; -  
 DR GO; GO:0004497; P:monooxygenase activity; IEA.  
 DR GO; GO:0006118; P:electron transport; IEA.  
 DR InterPro; IPR001128; Cytochrome\_P450.  
 DR Pfam; PF00067; P450; 1.  
 DR PRINTS; PR00385; P450.  
 DR PROSITE; PS00086; CYTOCHROME\_P450; 1.  
 KW Here; Monooxygenase; Oxidoreductase.  
 SQ SEQUENCE 525 AA; 59601 MW; B729F87AB214C1B9 CRC64;

Query Match 42.0%; Score 1147.5; DB 10; Length 525;  
 Best Local Similarity 43.7%; Pred. No. 3.2e-86;  
 Matches 229; Conservative 95; Mismatches 179; Indels 21; Gaps 7;

QY 4 ESSSWFIPKVLVLSVILSVIVKXSLMWRPKIEHFSKQIRGPPYHFIQNKELY 63  
 Db 11 EASPNLACAAAMAVLWLA-AWILEMAWMTPRRLGRALAEQIGKGTRYLFTGDVDENA 69

QY 64 GMLTKASSHPMPF-SHNILPRVLSFYHHMKRYGATFLVWGPPFRLTVADPDILREIFS 122  
 Db 70 RANKEARSKPLPLGSHDIIIPVQPMISNAIKENGLSTFWGPEPRVTIILDPESVREILS 129  
 QY 123 -KSEFEYKNEAPLVKQLEGGDILSLKGEKMAHRRKIISPFFHEMLKLVLPVYLKSVTD 181  
 Db 130 NKGHGXPRSSRFGLL-ANGLVNHQGEKMAKRRILNPAPFHEKIKRMLPVFSACSEE 188  
 QY 182 MYDKMSDKLSENGEVEVDVYEWQILTEDVTSRTAFSGSYEDGAVFRLQAQMLLCAEA 241  
 Db 189 MITRWNMSQGVSEVDVWPEFQULTGDIVISRTAFSGSYOEGTKIFOLQOGBRLMOA 248  
 QY 242 FQKVPIPGVFPFPTGNLKSRLDKERKSLKLIERRONALIDEGECKEPAKDLG 301  
 Db 249 FQTLFPGVFLPTKNRRMRREIDREICTILRGILIEKORAIKSGEA-----SSDDLIG 302  
 QY 302 LMIQAK-----NTVODIVECKSPFFAGKOTSNLLTWTIILSNHPEWQAKAR 351  
 Db 303 LLBSNRRESNGKADLGSTEDLIECKLFYFAGMETTSVLLTWTLLVLSHWPWQEQAR 362  
 QY 352 DEVLAVCGSDVPTKOHVVKLITSLMINESLRYPPIVATIRAKSDVKLGKXIPCGT 411  
 Db 363 KEVLHFFG-RTPDPENLSRLKIYVWLYEVLRLPRAIFVTRRYKAMELGITTPAGV 421  
 QY 412 ELIPIAVHDOAIMGNDVNEFPARFADGVPRAKHPVGFIPFGLGVRTICIGNTAIL 471  
 Db 422 NMLPILFIHDPNWKDASEFPORFADGISNAVXHPAAFPFGGPRICIGONFALL 481  
 QY 472 QAKTLAVMIQRTFLAPTYQAPVLMILYPOGATITRRRL 515  
 Db 482 EAKMALSTILQRFSELSPSYTHAPYVTLHPQGAIVLRKI 525

RESULT 10

Q9ATU4 PRELIMINARY; PRT; 525 AA.  
 ID Q9ATU4  
 AC Q9ATU4  
 DT 01-JUN-2001 (TREMBlrel. 17, Created)  
 DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)  
 DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)  
 DE Putative cytochrome P450.  
 OS Lolium rigidum.  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae;  
 OC Poaceae; Lolium.  
 OC NCBI\_TaxID=89674;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=SR\_31;  
 RA Fischer T.C.; Klattig J.T.; Gierl A.;  
 RT "A general cloning strategy of divergent plant cytochrome P450 genes  
 and its application in Lolium rigidum and Octium basilicum";  
 RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.  
 CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.  
 DR EMBL; AF321870; AAK38094.1; -  
 DR GO; GO:0004497; P:monooxygenase activity; IEA.  
 DR GO; GO:0006118; P:electron transport; IEA.  
 DR InterPro; IPR001128; Cytochrome\_P450.  
 DR Pfam; PF00067; P450; 1.  
 DR PRINTS; PR00385; P450.  
 DR PROSITE; PS00086; CYTOCHROME\_P450; 1.  
 KW Here; Monooxygenase; Oxidoreductase.  
 SQ SEQUENCE 525 AA; 59894 MW; D202EC96A3946F14 CRC64;

Query Match 42.0%; Score 1145.5; DB 10; Length 525;  
 Best Local Similarity 43.5%; Pred. No. 4.7e-86;  
 Matches 230; Conservative 96; Mismatches 178; Indels 25; Gaps 8;

QY 1 MEESSSMFIKVLVLSVILSVIVKXSLMWRPKIEHFSKQIRGPPYHFIQNKELY 58  
 Db 8 MLREASRSLAGAAAALMLLMAWYE--WAWMTPRRLRALQAQGRGQYRLFTGD 64  
 QY 59 VKELVGMKLASSHPMPF-SHNILPRVLSFYHHMKRYGATFLVWGPPFRLTVADPDIL 117

```

Db      65 VFNPAALNREARSKPPLGSHDI IQRVQMFENNVAKENGKRAFWFGTPTPRVMI PDELV 124
Qy      118 RELFS-KSEFEYKNEAHPYKQLEGDGLSLKGEKMAHKKIISPFTHEMLKILVYVL 176
Db      125 REVLNFKFGHYGKXSRIGKLL-ANGLANHQGEKMAHRRILNPAFHNEKIKMLVFA 183
Qy      177 KSVTDWVKMSDKLSENGEVEVDVYEMFQILTEVDISRTAFSSYEDGRAVFLQAQML 236
Db      184 TCCEMITRWDNMSQGSSEIDIMEFQNLGVISRTAFSSYQSGMKI FQJQSLAE 243
Qy      237 LCHEAFQKVFIPGRFPFRGNLKSRLDKIRKSLKILERRQNALIDGEGCEKPA 296
Db      244 RLIMAFQTFIPQYFWLPTKNNKRMALDCEIRILRGI ICKKXAKNGEA-----IS 297
Qy      297 KDLGLMIQAK-----NVTVDIVECKSFPAKQTSNLTWTITLISMHEW 346
Db      298 DDLGLLESNMRSENGKADLMSTEEIIQECKLYFAGMETTSVLLTTLISMHEW 357
Qy      347 QAKARDBVLRVCGSDVPTKDHVYKLTLSITLNESSLPLPYVATIRRAKSDVYKIGYK 406
Db      358 QEAKRDBVLRHG-RTTPDFEHLIRLKTVMILYEVRLPYPIITLIRRYKAMELGIGIK 416
Qy      407 IPCGTELLPIIAVHDDAIMGNDVNEFNPAFADGVPAKAFVGFIPGLGVRTICIG 466
Db      417 YPAGVNLMLPIFIHHDNMGKASERNPFRADGINSNAKHGSGFFPGGPRICIG 476
Qy      467 NLAIDKLTAVMIQRTFTFLAPTYQAPVYMLLYPGHAPITPRRL 515
Db      477 NFALBAKALSTIIQHFELBSPSYTHAPVYVITLHPQGAQIRIKI 525

```

## RESULT 11

```

Q9ATU5      PRELIMINARY;      PRT;      525 AA.
ID      Q9ATU5
AC      Q9ATU5;
DT      01-JUN-2001 (TReMBLrel. 17, Created)
DT      01-JUN-2001 (TReMBLrel. 17, Last sequence update)
DT      01-JUN-2003 (TReMBLrel. 24, Last annotation update)
DE      Putative cytochrome P450.
OS      Lolium rigidum.
OC      Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC      Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC      Poaceae; Lolium.
OC      NCBI_TaxID=89674;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=SLR 31;
RA      Fischer T.C., Klatzig J.T., Gierl A.;
RT      "A general cloning strategy of divergent plant cytochrome P450 genes
RT      and its application in Lolium rigidum and Oculum basilicum.";
RL      Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
CC      -1- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
DR      EMBL; AF21866; AAK3890.1; -.
DR      GO; GO:0004497; F:monooxygenase activity; IEA.
DR      GO; GO:0006118; P:electron transport; IEA.
DR      InterPro; IPR001128; Cytochrome_P450.
DR      Pfam; PF00067; P450; 1.
DR      PRINTS; PR00385; P450.
DR      PROSITE; PS00086; CYTOCHROME_P450; 1.
KW      Heme; Monooxygenase; Oxidoreductase.
SQ      SEQUENCE 525 AA; 59579 MW; 769A971E61064C42 CRC64;

```

Query Match 41.7%; Score 1138.5; DB 10; Length 525;  
 Best Local Similarity 43.3%; Pred. No. 1.8e-85;  
 Matches 227; Conservative 97; Mismatches 179; Indels 21; Gaps 7;

```

Qy      4 BSSGAPTEKVLVSLVLYKMSLMMRPKEHFSQGRGPPYHFFGNYKELY 63
Db      11 EASPMNLACAAAAMLLWLA-AMLEWMMMPRRGRALAEQKGTGRYLFQGDVPENA 69
Qy      64 GMLLKASHMPFP-SHNLPRVLSFYHWRKIYGTFLWNGPFRRLTVADPDILREIFS 122

```

```

Db      70 RANKEARSKPLPLGSHDIIPRVQPMISNAIKENGKLSFTWGPPEPRTVTLDPESVREILS 129
Qy      123 -KSEFEYKNEAHPYKQLEGDGLSLKGEKMAHKKIISPFTHEMLKILVYVLKSYTD 181
Db      130 NKFGHYKRPSSRFGKLL-ANGLVNQGKMAHRRILNPAFHNEKIKMLVFA 188
Qy      182 MVDKMSDKLSENGEVEVDVYEMFQILTEVDISRTAFSSYEDGRAVFLQAQMLCAEA 241
Db      189 MITWENMSQGSSEIDIMEFQNLGVISRTAFSSYQSGMKI FQJQSLAE 248
Qy      242 FQKVFIPGRFPFRGNLKSRLDKIRKSLKILERRQNALIDGEGCEKPAKDLIG 301
Db      249 FQTLFIQYFWLPTKNNRRAIDREI CTTLEGIIEKQRAKSGEA-----SSDILG 302
Qy      302 LMIQAK-----NVTVDIVECKSFPAKQTSNLTWTITLISMHEWQAKAR 351
Db      303 LLESNMRSENGKANGMSTEDIIIECKLYFAGMETTSVLLTTLISMHEWQOQAR 362
Qy      352 DEVLRVCGSDVPTKDHVYKLTLSITLNESSLPLPYVATIRRAKSDVYKIGYKIPGCT 411
Db      363 KEVLNHRG-RTTPDFEHLIRLKTVMILYEVRLPYPIITLIRRYKAMELGIGITYPAGV 421
Qy      412 ELIPIIAVHDDAIMGNDVNEFNPAFADGVPAKAFVGFIPGLGVRTICIGONTAL 471
Db      422 NLMPLFIHHDNMGKASERNPFRADGINSNAKHGSGFFPGGPRICIGNFALL 481
Qy      472 QAKTLTAVMIQRTFTFLAPTYQAPVYMLLYPGHAPITPRRL 515
Db      482 EAKMALSTIIQHFELBSPSYTHAPVYVITLHPQGAQIRIKI 525

```

## RESULT 12

```

Q9AX23      PRELIMINARY;      PRT;      519 AA.
ID      Q9AX23
AC      Q9AX23;
DT      01-JUN-2001 (TReMBLrel. 17, Created)
DT      01-JUN-2001 (TReMBLrel. 17, Last sequence update)
DT      01-JUN-2003 (TReMBLrel. 24, Last annotation update)
DE      Putative cytochrome P450.
GN      P045A01.12.
OS      Oryza sativa (Rice).
OC      Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC      Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC      Euphorbiaceae; Oryzae; Oryza.
OC      NCBI_TaxID=4530;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=cv. Nipponbare;
RA      Sasaki T., Matsumoto T., Yamamoto K.;
RT      "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 1, PAC
RT      clone: P045A01.1";
RL      Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
CC      -1- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
DR      EMBL; AF002899; BAB21156.1; -.
DR      Gramene; Q9AX23; -.
DR      GO; GO:0004497; F:monooxygenase activity; IEA.
DR      GO; GO:0006118; P:electron transport; IEA.
DR      InterPro; IPR001128; Cytochrome_P450.
DR      Pfam; PF00067; P450; 1.
DR      PRINTS; PR00385; P450.
KW      Heme; Monooxygenase; Oxidoreductase.
SQ      SEQUENCE 519 AA; 59124 MW; DA48BEE0FEF7B4 CRC64;

```

Query Match 41.2%; Score 1126; DB 10; Length 519;  
 Best Local Similarity 44.0%; Pred. No. 1.9e-84;  
 Matches 218; Conservative 102; Mismatches 156; Indels 20; Gaps 6;

```

Qy      32 WMRPKIEHFSQGRGPPYHFFGNYKELYGMLKASHMPFP-SHNLPRVLSFYH 90
Db      32 WMRPKIEHFSQGRGPPYHFFGNYKELYGMLKASHMPFP-SHNLPRVLSFYH 91
Qy      91 WKIKYATFLWNGPFRRLTVADPDILREIFS-KSEFEYKNEAHPYKQLEGDGLSLK 149

```

Mon May 24 08:10:11 2004

us-09-992-901-2.rspt

Page 8

```
DB 92 VLKHKLSFWMTGPKPVILTRDPLAREILSNKSNFAKQTAGIAKFWG-GVVTYEG 150
QY 150 EKWMAHRIISPTHEMELKLVAVLKSITDWDKSDLSNENVEDVYEMOILTE 209
DB 151 EKWMAHRIILPAFOEKIKRMVLVLAACCKITRWVSNSESGISELDVDEQNLIG 210
QY 210 DVISRTAFSSYEDGRAVFLQAOQMLCAEAFQKVFIPGYRFPPTGNLKSRLDKER 269
DB 211 DVISRTAFSSYEDGRAVFLQAOQMLCAEAFQKVFIPGYRFPPTGNLKSRLDKER 270
QY 270 KSLILERRRQNMIDGEGECKEPAKDLGLMIQAK-----NVTQDIVECK 319
DB 271 TLKRIIVKRDKAVRNGSGN-----DDELGLVESNRSQSEKEDVMSJEDITECK 324
QY 320 SEFFAGKQTSNLLTWTITLISMEPEMOQARDEVLRVCGSDVPTRKHVVKLTKLSMIL 379
DB 325 LFYAAGSEITSMILTWITLILSMHPEWQOAREEVNHFQ-RTTPDHDLISRLKIVTIL 383
QY 380 NESLTPYVATIRASDVKLGKYPICGTETLLPILAVHDDQIWGNDVNEPAPF 439
DB 384 HEVLRIPVVFLOQTTHKELEGGIYFEGVNFILPVLSIHDSIMQDAIKENPERF 443
QY 440 ADGVRPAKHEVGFIPFGLGVRTIGONLAILQAKTLAVMTQRTFLAPTYOAPTVL 499
DB 444 ANGVSKATKFTOTAFESFAMGPRICLGOSFALIEAKALATILQSFSELSPYTHAPTV 503
QY 500 MLVYPOHGAPTFRRL 515
DB 504 LILPOQVGSPIKMKL 519

RESULT 13
Q9LUD2 PRELIMINARY; PRT; 515 AA.
AC Q9LUD2;
DT 01-OCT-2000 (TREMblrel. 15, Created)
DT 01-OCT-2000 (TREMblrel. 15, Last sequence update)
DT 01-JUN-2003 (TREMblrel. 24, Last annotation update)
DE Cytochrome P450 (A13G14620/MIE1.12).
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eucosids II; Brassicales; Brassicaceae; Arabidopsis.
ON NCBI_TaxId=3702;
RX SEQUENCE FROM N.A.
RP STRAIN=Columbia;
RA Sato S., Nakamura Y., Kaneko T., Kato T., Asamizu E., Tabata S.;
RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RC SEQUENCE FROM N.A.
RX MEDLINE=20277480; PubMed=10819329;
RA Nakamura Y.;
RT "Structural analysis of Arabidopsis thaliana chromosome 3. I. Sequence
RT features of the regions of 4,504,864 bp covered by sixty P1 and TAC
RT clones."
RL DNA Res. 7:131-135(2000).
RN [3]
RP SEQUENCE FROM N.A.
RA Cheuk R., Chen H., Kim C.J., Koesema E., Meyers M.C., Banu J.,
RA Bower L., Carninci P., Dale J.M., Goldsmith A.D., Hayashizaki Y.,
RA Ishida J., Jiang P.X., Jones T., Kamiya A., Karlin-Neumann G.,
RA Kawai J., Lam B., Lee J.M., Lin J., Liu S.X., Miranda M., Natuska M.,
RA Nguyen M., Onodera C.S., Palm C.J., Pham P.K., Quach H.L., Sakurai T.,
RA Saitou M., Seki M., Southwick A., Tang C.C., Toriumi M., Yamada K.,
RA Yamamura Y., Yu G., Yu S., Shinozaki K., Davis R.W., Theologis A.,
RA Ecker J.R.;
RT "Arabidopsis cDNA clones."
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
CC -1 SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
DR EMBL; AB023038; BAB02394.1; -
DR EMBL; AY052208; AAK97679.1; -
```

```
DR GO; GO:0004497; P:monooxygenase activity; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR001128; Cytochrome_P450.
DR Pfam; PF000677; P450_1.
DR PRINTS; PR00385; P450.
DR PROSITE; PS00086; CYTOCHROME_P450_1.
DR Heme; Monooxygenase; Oxidoreductase.
SQ SEQUENCE 515 AA; 58643 MW; 9F4BFCF7686F55A1 CRC64;

Query Match 41.3%; Score 1124.5; DB 10; Length 515;
Best local similarity 40.3%; Pred. No. 2,5e-84;
Matches 209; Conservative 126; Mismatches 162; Indels 21; Gaps 6;

QY 12 KYLVLSVILSLVYV-----KMSLLMRPKRIEHSFKQIGRPYHFIQNYKELV 63
DB 5 KISAVAVAAVVVYVTVVIMKGLNVAMLRPKNEAYLRQGLSGTPFVLGDIKRA 64
QY 64 GMLKASSHPPESHNLPVLSFYHMKIYGFVLPWPGPTFRLLTVADPLREIFSK 123
DB 65 SWVEDEKSPINLDDYTHRMPLIQOTYKDKHGSYMMGPASVIVTKPHIKDVNR 124
QY 124 SEFYERKNAHPIVQLBEGDGLSLKGRKMAHKKIISPTHEMNLKLVAVLKSVTDNV 183
DB 125 VYDFPKPVHPIV-ELFATGVALYGEKWSKHKIIPSFHLEKIMIPAFYESCEMI 183
QY 184 DKMSDKLSNGEV-EVDYEWFOILTEDVLSRTAFSSYEDGRAVFLQAOQMLCAEAF 242
DB 184 SKMEKLVTEQSSNEIDWPIGLDLSVSRNAGSSYEGKRIELDEQGRVLYKL 243
QY 243 QYVFIQYRFPPTGNLKSRLDKERISLLKILERRQNALDGESECKEPAKDLGL 302
DB 244 ELAFIPGKRFPLTKNLRKQINKEVSRREIIMKQQRMDGKAPK-----NDLIGI 297
QY 303 MIOAKN-----YVQDIVECKSPFPAGKQTSNLLTWTITLISMEPEMOQARDEVLRVC 358
DB 298 LIESNSGHSISIEDVEECCRLFPAGQETVALVLTWIMLSHQKMDQAREILKXI 357
QY 359 GSRDVPYDHYVVKLTKTISMLINSLRYPPIVATIRASDVKG-GYKIPCGTELLPT 417
DB 358 GGNRPKNDALSRKTKSMILNEVLRILYPPGILGRVETELGSDMTLPGAQVAVPV 417
QY 418 IAVHDDQIWMGNDVNEPAPFADGVRPAKHEVGFIPFGLGVRTIGONLAILQAKTL 477
DB 419 LNVHDDPELVKEDVHEENPERFADGISKARKQVSLFPCWGRFPDQGNFALMEKXAL 477
QY 478 AVMIQRTFLAPTYOAPTVLMLVYPOHGAPTFRRL 515
DB 478 VLIQRPSEFELSPSYTHAPTVLTHPOFGAPLIFMVL 515

RESULT 14
Q9WLE1 PRELIMINARY; PRT; 515 AA.
AC Q9WLE1;
DT 01-MAR-2002 (TREMblrel. 20, Created)
DT 01-MAR-2002 (TREMblrel. 20, Last sequence update)
DT 01-JUN-2003 (TREMblrel. 24, Last annotation update)
DE AT3G14620/MIE1.12.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eucosids II; Brassicales; Brassicaceae; Arabidopsis.
ON NCBI_TaxId=3702;
RX SEQUENCE FROM N.A.
RP STRAIN=Columbia;
RA Kim C.J., Chen H., Cheuk R., Meyers M.C., Banu J., Bower L.,
RA Carninci P., Chang E., Dale J.M., Goldsmith A.D., Hayashizaki Y.,
RA Ishida J., Jones T., Kamiya A., Karlin-Neumann G., Kawai J., Lam B.,
RA Lee J.M., Lin J., Miranda M., Natuska M., Nguyen M., Onodera C.S.,
RA Palm C.J., Quach H.L., Sakurai T., Saitou M., Seki M., Southwick A.,
RA Tang C.C., Toriumi M., Yu H.C., Yamada K., Yamamura Y., Yu G., Yu S.,
RA Shinozaki K., Davis R.W., Theologis A., Ecker J.R.;
RT "Arabidopsis ORF clones."
DR EMBL; AB023038; BAB02394.1; -
DR EMBL; AY052208; AAK97679.1; -
```

RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.  
 CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.  
 DR EMBL; AF46870; AL48603.1; -  
 DR GO; GO:0004497; F:monooxygenase activity; IEA.  
 DR GO; GO:0006118; P:electron transport; IEA.  
 DR InterPro; IPR001128; Cytochrome\_P450.  
 DR Pfam; PF00067; P450.1.  
 DR PRINTS; PR00385; P450.  
 DR PROSITE; PS00086; CYTOCHROME\_P450.1.  
 DR Heme; Monooxygenase; Oxidoreductase.  
 SQ SEQUENCE 515 AA; 5861 MW; 9FAAD6C768648CA1 CRC64;

Query Match 41.2%; Score 1124.5; DB 10; Length 515;  
 Best Local Similarity 40.3%; Pred. No. 2.5e-84;  
 Matches 209; Conservative 126; Mismatches 162; Indels 21; Gaps 6;

QY 12 KVLVLSVLSLVY-----KMSLLMRPRKIEHFKQIGRGPYHFFIGNVKELY 63  
 DB 5 KISAVAVAAVVVTVTWIMKGLNVAALPKNEAYLKRQGLSGTFEFLVDIKREA 64  
 QY 64 GMMKASHPMPFSNHLPRVLSFYHMRKIYGATFLWFGPTFRLTVADPDIREIFSK 123  
 DB 65 SIVDEKSRPINTLDYHRRVPLIQGVKHGKTSYMMGPIASVITKEBHIDVLR 124  
 QY 124 SEFYKNEAHPVVKQLEGDGLSLKGEKMAHRRKIISPTFMENLKLVVVLKSVTDVY 183  
 DB 125 YVDFEKPVPVHIV-ELFATGVALYEGEKMSGRKIIIPSPFLEKIKIMPAFYESGSEMI 183  
 QY 184 DKWSKLSENGEV-EVDYVWFQLTTEVTSRTFGSSYSDGRAVFLQAQMLCAEAF 242  
 DB 184 SWEKLVLEQSSNIDWPFYDGLDTSRTIFGSSYERKRIFFLOEBOGRVYLKL 243  
 QY 243 QKVFYGRFFPTFRGNLSRKLDKEIRKSLIKLIERRONALIDEGECKEPAKADILGL 302  
 DB 244 ELAFIPGRFLPTKSNLMRQINKEVKSRLREIMKRGMDTEAKR-----NDLLGI 297  
 QY 303 MIAQAK-----VTVQDIVECKSFPAAGQQTNSLLTWTLLSHPEMOKARDEVLRVC 358  
 DB 298 ILSESGDHGMSIEVVEECRLPHFAGQETAVLWMTLSHQQMQUAEEILKVI 357  
 QY 359 GSRDVPYKDVVKLTSLMILNESLRLPPVATIRAKSDVKLG-GYKIPCGTELLIPT 417  
 DB 358 GKNNKPNPDALSRITMNLNEVRLKPPGILGRVYEKTKGEMTLFGAQVYIPV 417  
 QY 418 IAVHDDQAIWQNDVNEFNPAPPAQGVPRPAAGVGFIPGLGVATCCGQNTALIQAKLT 477  
 DB 418 LNVHDDPELMGEDVHEFPERPADGISKATKNQVSFLPFQWGFPCGQNFALMEAKAL 477  
 QY 478 AVMIORFPHLAPYQHAATVMLLYPOHGAIPITPRL 515  
 DB 478 VTILOKSPFELSPSTTHAPHTVLTILHPOGAPLIPHML 515

## RESULT 15

Q8LGN4 PRELIMINARY; PRT; 528 AA.

AC Q8LGN4; 01-OCT-2002 (T-EMBLrel. 22, Created)  
 DT 01-OCT-2002 (T-EMBLrel. 22, Last sequence update)  
 DT 01-JUN-2003 (T-EMBLrel. 24, Last annotation update)  
 DE Cytochrome P450 monooxygenase CYP72A26.  
 OS Zea mays subsp. mays (maize).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 OC PACAD clade; Panicoideae; Andropogonaceae; Zea.  
 OC NCBI\_taxid=4578;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. Inbred B73;  
 RA Wang J., Schuler M.;  
 RT "Characteristics and Transcriptional Regulation of Maize Cytochrome  
 P450 Monooxygenases.";  
 RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.

RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. Inbred B73;  
 RA Wang J., Reed R., Schuler M.;  
 RT "Xenobiotics Regulate Maize Cytochrome P450 Monooxygenases at a Post-  
 Transcriptional Level.";  
 RT Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.  
 CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.  
 DR EMBL; AY01866; AL60592.1; -  
 DR GO; GO:0004497; F:monooxygenase activity; IEA.  
 DR GO; GO:0006118; P:electron transport; IEA.  
 DR InterPro; IPR001128; Cytochrome\_P450.  
 DR Pfam; PF00067; P450.1.  
 DR PRINTS; PR00385; P450.  
 DR PROSITE; PS00086; CYTOCHROME\_P450.1.  
 DR Heme; Monooxygenase; Oxidoreductase.  
 SQ SEQUENCE 528 AA; 59972 MW; BAIED8D207E6B60B CRC64;

Query Match 41.0%; Score 1119; DB 10; Length 528;  
 Best Local Similarity 43.1%; Pred. No. 7.4e-84;  
 Matches 214; Conservative 103; Mismatches 164; Indels 16; Gaps 7;

QY 28 MSLLMRPRKIEHFKQIGRGPYHFFIGNVKELYGMKASHPMPF-SNHLPRVLS 86  
 DB 39 LRLMQPRRLERALRAQVKGTSYFPADDLKYEGRSLAKEAWSKPLPRCHDTARVAP 98  
 QY 87 FYHMRKIYATFLWFGPTFRLTVADPDIREIFS-KSEFYKNEAHPVVKQLEGDGL 145  
 DB 99 FLHRLVLEHGTSWFGSPKVTIVDPELAKOVLNKGHFEKLVKPAISKYL-GSGVA 157  
 QY 146 SLKGEKMAHRRKIISPTFMENLKLVVVLKSVTDVDMKW-SDKLSENGEVVDYVWF 204  
 DB 158 SHGEKMAHRRKIISPTFMENLKLVVVLKSVTDVDMKW-SDKLSENGEVVDYVWF 217  
 QY 205 QLTEDVTSRTAFSSYEDGRAVFRLOAQOMLCAEFQKV-FIPGRRPPTFRGNLSRK 263  
 DB 218 QNLTGDVTSRTAFSSYEDGRAVFRLOAQOMLCAEFQKV-FIPGRRPPTFRGNLSRK 277  
 QY 264 LDKERKSLTLIRRRONALIDEGECKEPAKADILGLMIOAK-----NTVQDIVEEC 318  
 DB 278 NNREVSLEIRLITKRIQAMERGQTK-----DMLGLLETNMRDDMGMTIEDVIEEC 331  
 QY 319 KSFFPAGKQTSNLTWTTLTLLSHPEMOKARDEVLRVCGSDVPTKDVVVLKTLISMT 378  
 DB 332 KVFYFAGMETTSVLLTWTMVLVLSHPEMOKARDEVLRVCGSDVPTKDVVVLKTLISMT 391  
 QY 379 LNESRLKPPVATIRAKSDVKLGKTKCGTELLIPTAVHDDQAIWQNDVNEFNPAP 438  
 DB 392 LVEVRLKPPVATIRAKSDVKLGKTKCGTELLIPTAVHDDQAIWQNDVNEFNPAP 451  
 QY 439 FADGVPRAAKHVGFIPGLGVATCCGQNTALIQAKLTAVMIORFPHLAPYQHAPTV 498  
 DB 452 FSDGVSRASKDPGAFLPFGWGPRIQGNFALLEAKMALCMIIQRFEPRLAPSTHAPHT 511  
 QY 499 LMLYPOHGAIPITPRL 515  
 DB 512 VTIHMPGDAQKLRAT 528

Search completed: May 19, 2004, 10:27:20  
 Job time : 49 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 19, 2004, 10:24:45 ; Search time 23 Seconds  
(without alignments)  
1167.196 Million cell updates/sec

Title: US-09-992-901-2

Perfect score: 2730  
Sequence: 1 MEESSSMFIPKVLVLSVIL.....LTPQHGAITFRRLNHD 520

Scoring table:

BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents AA:\*  
1: /cgm2\_6/ptodata/2/iaa/5A.COMB.pep.\*  
2: /cgm2\_6/ptodata/2/iaa/5B.COMB.pep.\*  
3: /cgm2\_6/ptodata/2/iaa/6A.COMB.pep.\*  
4: /cgm2\_6/ptodata/2/iaa/6B.COMB.pep.\*  
5: /cgm2\_6/ptodata/2/iaa/PCTUS.COMB.pep.\*  
6: /cgm2\_6/ptodata/2/iaa/backfile1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2730	100.0	520	4 US-09-527-073-2	Sequence 2, Appl
2	482.5	17.7	489	4 US-09-852-067-4	Sequence 4, Appl
3	479.5	17.6	524	4 US-09-976-594-533	Sequence 533, App
4	473	17.3	510	4 US-09-852-067-2	Sequence 2, Appl
5	441.5	16.2	504	4 US-08-457-274A-25	Sequence 25, Appl
6	441.5	16.2	504	5 PCT-US95-05758-25	Sequence 25, Appl
7	418	15.3	503	4 US-09-144-367-2	Sequence 2, Appl
8	404	14.8	576	3 US-08-948-564-16	Sequence 16, Appl
9	386	14.1	503	4 US-09-583-447A-2	Sequence 2, Appl
10	377.5	13.8	504	4 US-09-583-447A-2	Sequence 4, Appl
11	367	13.4	557	4 US-09-518-386B-1	Sequence 1, Appl
12	367	13.4	557	4 US-09-518-386B-3	Sequence 3, Appl
13	347	12.7	502	4 US-09-904-615-69	Sequence 69, Appl
14	344.5	12.6	507	4 US-08-457-274A-23	Sequence 23, Appl
15	344.5	12.6	507	5 PCT-US95-05758-23	Sequence 23, Appl
16	335.5	12.3	511	3 US-08-991-677-4	Sequence 4, Appl
17	331.5	12.1	507	1 US-08-457-274A-22	Sequence 22, Appl
18	331.5	12.1	507	5 PCT-US95-05758-22	Sequence 22, Appl
19	329.5	12.1	496	3 US-09-292-768-66	Sequence 66, Appl
20	329.5	12.1	517	4 US-09-302-620B-100	Sequence 100, App
21	329.5	12.1	517	4 US-09-911-781-32	Sequence 32, Appl
22	328.5	12.0	496	3 US-09-292-768-64	Sequence 64, Appl
23	327.5	12.0	496	3 US-08-881-784-1	Sequence 1, Appl
24	327.5	12.0	496	3 US-09-292-768-2	Sequence 2, Appl
25	327.5	12.0	496	3 US-09-172-339-6	Sequence 6, Appl
26	323	11.8	520	4 US-09-564-808-4	Sequence 4, Appl
27	323	11.8	540	4 US-09-302-620B-98	Sequence 98, Appl

28	320.5	11.7	497	3 US-08-724-466B-4	Sequence 4, Appl
29	320.5	11.7	497	4 US-08-882-164D-4	Sequence 4, Appl
30	318	11.6	508	4 US-09-126-420A-25	Sequence 25, Appl
31	318	11.6	540	4 US-09-302-620B-99	Sequence 99, Appl
32	316.5	11.6	420	4 US-09-583-447A-6	Sequence 6, Appl
33	316	11.6	476	1 US-08-313-075A-30	Sequence 30, Appl
34	315	11.5	523	3 US-09-302-620B-95	Sequence 95, Appl
35	314.5	11.5	497	4 US-08-882-164D-32	Sequence 32, Appl
36	314.5	11.5	509	3 US-08-948-564-18	Sequence 18, Appl
37	314.5	11.5	517	4 US-09-302-620B-101	Sequence 101, App
38	314	11.5	199	4 US-09-351-229-2	Sequence 2, Appl
39	312.5	11.4	508	3 US-08-991-677-2	Sequence 2, Appl
40	311	11.4	520	4 US-09-564-808-2	Sequence 2, Appl
41	308	11.3	529	4 US-09-615-192A-405	Sequence 405, App
42	305.5	11.2	513	3 US-08-948-564-2	Sequence 2, Appl
43	305	11.2	509	4 US-09-499-302A-8	Sequence 8, Appl
44	305	11.2	520	2 US-09-091-432-2	Sequence 2, Appl
45	305	11.2	520	4 US-09-387-663-2	Sequence 2, Appl

#### ALIGNMENTS

```
RESULT 1
US-09-527-073-2
; Sequence 2, Application US/09527073
; Patent No. 6534313
; GENERAL INFORMATION:
; APPLICANT: Michael M. Neff
; APPLICANT: Joanne Chory
; TITLE OF INVENTION: GENETICALLY MODIFIED PLANTS HAVING
; TITLE OF INVENTION: MODULATED BRASSINOSTEROID SIGNALING
; FILE REFERENCE: SALKINS.024A
; CURRENT FILING DATE: US/09/527,073
; PRIOR APPLICATION NUMBER: 2000-03-16
; PRIOR FILING DATE: 1999-03-16
; PRIOR APPLICATION NUMBER: US 60/124570
; PRIOR FILING DATE: 1999-03-16
; PRIOR APPLICATION NUMBER: US 60/170,931
; PRIOR FILING DATE: 1999-12-14
; PRIOR APPLICATION NUMBER: US 60/172,832
; PRIOR FILING DATE: 1999-12-20
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 520
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-09-527-073-2

Query Match      100.0%; Score 2730; DB 4; Length 520;
Best Local Similarity 100.0%; Pred. No. 4.1e-278;
Matches 520; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 MEESSSMFIPKVLVLSVILSVYVGMGLMRRPKIIEHFEKQIGRPPIFFGNK 60
1 MEESSSMFIPKVLVLSVILSVYVGMGLMRRPKIIEHFEKQIGRPPIFFGNK 60
61 ELVGMMLKASHPVPSHNTLPRVLSFYHMRKIYATFLVWGPFTFLVADPDILREI 120
61 ELVGMMLKASHPVPSHNTLPRVLSFYHMRKIYATFLVWGPFTFLVADPDILREI 120
61 ELVGMMLKASHPVPSHNTLPRVLSFYHMRKIYATFLVWGPFTFLVADPDILREI 120
61 ELVGMMLKASHPVPSHNTLPRVLSFYHMRKIYATFLVWGPFTFLVADPDILREI 120
121 FSKSEFYEKNEAPLVQLEGDGLSLKGEKMAHREKTIPTFHMENTKLIVVYVYKSTV 180
121 FSKSEFYEKNEAPLVQLEGDGLSLKGEKMAHREKTIPTFHMENTKLIVVYVYKSTV 180
121 FSKSEFYEKNEAPLVQLEGDGLSLKGEKMAHREKTIPTFHMENTKLIVVYVYKSTV 180
121 FSKSEFYEKNEAPLVQLEGDGLSLKGEKMAHREKTIPTFHMENTKLIVVYVYKSTV 180
181 DMVDKMSKSENEVEVDVYEWQILTEVDVISTARSSYEDGRAVFRLOAOQMLCAE 240
181 DMVDKMSKSENEVEVDVYEWQILTEVDVISTARSSYEDGRAVFRLOAOQMLCAE 240
181 DMVDKMSKSENEVEVDVYEWQILTEVDVISTARSSYEDGRAVFRLOAOQMLCAE 240
181 DMVDKMSKSENEVEVDVYEWQILTEVDVISTARSSYEDGRAVFRLOAOQMLCAE 240
241 AFQKVFIPGYVFPFTFGNLKSRKLDKEIRKSLKTLERRONAIDGEGECKEPAPKDLL 300
241 AFQKVFIPGYVFPFTFGNLKSRKLDKEIRKSLKTLERRONAIDGEGECKEPAPKDLL 300
241 AFQKVFIPGYVFPFTFGNLKSRKLDKEIRKSLKTLERRONAIDGEGECKEPAPKDLL 300
```

Mon May 24 08:10:10 2004

us-09-992-901-2.rai

Page 2

Qy	301	GLMIQAKNTAVDIDYEECKSPFFAGQTTSNLLTITLLSMHEWQAKRDEVLRCGS	360
Qy	302	GLMIQAKNTAVDIDYEECKSPFFAGQTTSNLLTITLLSMHEWQAKRDEVLRCGS	360
Db	301	GLMIQAKNTAVDIDYEECKSPFFAGQTTSNLLTITLLSMHEWQAKRDEVLRCGS	360
Qy	361	RDVPTKQHVVKLKTLSMLINESLRYLPVATIRPAKSDVKLGSKYIPCGTELLIPIAV	420
Db	361	RDVPTKQHVVKLKTLSMLINESLRYLPVATIRPAKSDVKLGSKYIPCGTELLIPIAV	420
Qy	421	HHQQAAGWDVNEFNARFADGVPRAAKHPVGFIPFGIGVTCIGQNALIQAULTAVM	480
Db	421	HHQQAAGWDVNEFNARFADGVPRAAKHPVGFIPFGIGVTCIGQNALIQAULTAVM	480
Qy	481	IQRFTHLAFTYQHAFTVMLLYPGHGAFTFRRLTINED	520
Db	481	IQRFTHLAFTYQHAFTVMLLYPGHGAFTFRRLTINED	520

```

1 RESULT 2
2 US-09-852-067-4
3 : Sequence 4, Application US/09852067
4 : Patent No. 6531287
5 : GENERAL INFORMATION:
6 : APPLICANT: MERKTOV, Gennady et al
7 : TITLE OF INVENTION: ISOLATED HUMAN DRUG-METABOLIZING
8 : TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES ENCODING HUMAN
9 : TITLE OF INVENTION: DRUG-METABOLIZING PROTEINS,
10 : TITLE OF INVENTION: AND USES THEREOF
11 : FILE REFERENCE: CLO08097-CIP
12 : CURRENT APPLICATION NUMBER: US/09/852,067
13 : CURRENT FILING DATE: 2001-05-10
14 : NUMBER OF SEQ ID NOS: 4
15 : SOFTWARE: FastSeq for Windows Version 4.0
16 : SEQ ID NO 4
17 : LENGTH: 469
18 : TYPE: PRT
19 : ORGANISM: Human
20 : US-09-852-067-4

```

Query Match	17.7%;	Score 482.5;	DB 4;	Length 489;
Best Local Similarity	27.1%;	Pred. No. 8.8e-42;		
Matches 141;	Conservative 103;	Mismatches 215;	Indels 61;	Gaps 15;

```

QY 2 LSVIAIKGNSL-----MMRPRKLEEHFSKQGLNGPPIHFFISNVVLWGMMLKASSHFF 16
QY 7 LLLILIKAAQLYHROMLBRALQ-----FCEPPFHMLGSHRE-----F 46
Db 77 SHNILEPRVLSYTHMRKTIYGATFLVMF--GPIRFLTVADPDLIRIFESKSEFFYKKNAMHL 135
QY 47 QND--QELERIQKWEKEPGACPMWLSGNKRALLVYDPDLKYLIGRSD--PRAPPNYKL 102
Db 136 VKQJEDGGLSLKGEKMAHHRKISPTFEMENLKLYPVVLKSYTDWVDMQSDKSENGE 195
QY 103 MTRWITGYGLLIDGQVWFOHRKMLTPAHYDILKPYGLWAVDSQVIMLDMBELLIQDSS 162
Db 196 VEVVYWEFQILTEDVISRTAF--GSSTYEDGRAVFRQAQOML--LCAEAFQVVF---- 246
QY 163 LE--IFQVSLMTLDTIMKCAFSGYQVDLRNHSYQALINDLNNLYFRANVHQSD 220
Db 247 -----IPGRPFPTRGNLKSRKLDSEIKSKLLKLERRQVADIGBEECKEPAARDILG 301
QY 221 FLVRLSBEGLFLPRACQLAHEHTDVIQ-----QRQAQIQGGELEKRYRRKRLDPLD 273
Db 302 LMIQAK-----NVTVODIVECKSEFFPAQKOTSNILTWITLLSKMPEVQAKARDEVLR 356
QY 274 VLLFAKENGSSLSDDDLAAVDTLMFSGHDTTASGVSWIYALATPRHQHCRREIQG 333
Db 357 VCGSSRVPTRKDVYVKLTKLSMTLINESLRYPPIYATTIRAKSDYKL--GGYKIPCGBELL 415
QY 334 ILDGDASTIWEHDQMPYTTMCIKALRTYPPVPSVTQSLKPYTFPDGSLPKGYILFL 393
Db 416 PIAVHHDDALINQDINVEENPARPAPDQVPRAKHVGIFPGGLVRFCTIGONALAIQAKL 475
QY 394 SIYGLHNPKYKQNP--EVFDPFRFA--PDSAKISHALPLPGSGARRCIGKQPMELKV 449

```

D5  
QY 476 TLAVNIQRFTHLAPTYGNAFTVLLLYQHGRITERRL 515  
          ::: |||         ::: :::: ::  
D6 450 AVATLLIRFEELLPDPTRVIPPIARVVLKSKNGIHLRLKL 489

RESULT 3  
US-09-976-594-533

```

1 Sequence 533, Application US/09976594
2 Patent No. 6673549
3 GENERAL INFORMATION:
4 APPLICANT: Furness, Michael
5 APPLICANT: Buchbinder, Jenny
6 TITLE OF INVENTION: GENES EXPRESSED IN C3A LIVER CELL CULTURES TREATED WITH STEROID
7 FILE REFERENCE: PA-0041 US
8 CURRENT APPLICATION NUMBER: US/09/976,594
9 CURRENT FILING DATE: 2001-10-12
10 PRIOR APPLICATION NUMBER: 60/240,409
11 PRIOR FILING DATE: 2000-10-12
12 NUMBER OF SEQ. ID NOS: 1143
13 SOFTWARE: PERL Program
14 SEQ ID NO 533
15 LENGTH: 524
16 TYPE: PRT
17 ORGANISM: Homo sapiens
18 FEATURE:
19 NAME/KEY: misc feature
20 OTHER INFORMATION: incyte ID No. 6673549 6825202CD1
21 US-09-976-594-533

```

Query Match	17.6%	Score 479.5	DB 4	Length 524
Best Local Similarity	28.4%	Pred. No. 2e-41		
Matches 154	Conservative 83	Mismatches 207	Indels 99	Gaps 21

```

QY      1 MEESSSNFIPKVLVSVILSVIVKMSLL-----WNRPRRIEEHHSKQIGRGPYHF 55
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      1 MPQLSLWLGIGPVAAASPMLILLVGGSWLLARVLAW-----TYTF 41

```

```

QY      56  IGVNKELVGNMFKASSHPMPFSNLLPRVLSVYHMKRI-----YGVLF  99
          | | | | | | | | | | | | | | | | | | | | | | | |
Db      42  YDNCRL-----QCFPQF-----PKQNMFWGHQGLVTPFEGMKTLTQLVTTYQGF  88
          | | | | | | | | | | | | | | | | | | | | | | | |

```

```

QY      100 LWFEGPIRILIV-ADPDLKRLKLSKSFYEKNEA--HPLVKLGGDGLSLNGENWNRK 13
      : ||||| : ||||| : : ||||| : ||||| :
Db      89 KMLGPIFPLILICHPIIRPTSASAAVAPKDMIFFGLKRWLGDGLLSCGDKNMSRRH 14

```

```

157 KIISPTEHMEKALVAVLVSVIDMVKMSJLSENGEVEVIVIEFQIUEEDVJSLIA 21
    ::: : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 149 RMUTPAFHENILKPYMKIFNKSVINIMDKM-QRLASGSAARDMFHISLMTLDSTQKV 20

```

```

21 / FSSIED-----KAVTLYA-----QILDCHEM VAVT LGXNL ANVAVC
      | : | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 208 F--SFEENCQEKPEYIAILBLSAFVEKRNQIILLTDFLYLTPDGQRF-----RRA 25

```

260 CHLVHDFDAVIQ-ERRRTLTQGIIDFLKNKAKSKTIDFIDVLLSKDEDEKELSDSD 311

318 IRBADTFMEFGHDTIASGLSNVLYHLAKPEYQOCRCQEVVELLKORE-PIEIMWDLA 37

```

377 QLEFLTMCKESRLHPEVPVVISRCCQDFVLBDGRIVPKGIVCLINIIIGHNPTWV-P 433

```

436 DPEVYDPRF-DQENIKERSPLAFIPSAGPNNCIGAFAMKEMKVIALTL--LHFRIL 43

Db	493 PTH 495
----	-------------

Mon May 24 08:10:10 2004

us-09-992-901-2.rat

Page 3

```

RESULT 4
US-09-852-067-2
; Sequence 2, Application US/09852067
; Patent No. 6531297
; GENERAL INFORMATION:
; APPLICANT: MERKULOV, Gennady et al
; TITLE OF INVENTION: ISOLATED HUMAN DRUG-METABOLIZING
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES ENCODING HUMAN
; TITLE OF INVENTION: DRUG-METABOLIZING PROTEINS,
; TITLE OF INVENTION: AND USES THEREOF
; FILE REFERENCE: C0000897-CIP
; CURRENT APPLICATION NUMBER: US/09/852,067
; CURRENT FILING DATE: 2001-05-10
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 510
; TYPE: PRT
; ORGANISM: Human
US-09-852-067-2

Query Match      17.3%; Score 473; DB 4; Length 510;
Best Local Similarity 25.7%; Pred. No. 9,4e-41;
Matches 138; Conservative 111; Mismatches 221; Indels 68; Gaps 19;

QY      8 WFIKRVLVSVILSLVVKGMSLIMWPKIEEHFSKQIRGPPYHFFIGNVKELVGMML 67
DB      10 WAPRYLAFFVCLALGLQAIKYLRRQRLRD---LRPPAPPTWPFQHQKFI----- 61
QY      68 KASHMPFSSNLLPVLSFPHWKRKYGATFLVPGPTR--LTVADPLLSEIRSKSE 125
DB      62 --ODNNWKELEIEIK-----YPPAPFPWIGP--FOAFCIYDPDYAKTILSRID 107
QY      126 ---FYEKNEAHPVVKQLEGDLISLKEKWAHHRKIIISPTFHEMNIKLVPVVLKSVTDM 182
DB      108 PKSRYLQKSPPL----GKSLALDGPKNQHRRLLPFGFHNILKAYEVWASGVKMM 163
QY      183 VDKSDDKLSNGEVEVDYEMFQILTEDVISRTAFG-----SSYED--GRAVRL--- 230
DB      164 LDKW-EKICSTQDSVEYEHINMSLDIIMKCAFSEKTCQNTSTHDPYAKIIFELSKI 222
QY      221 ---QAQOMLCAEAFQKFIPIGYRFPFRGSLKSKDKIERSLKLIERRQNAIDSE 287
DB      223 IFRHRYSLYHSDIIFKLSFGYRF-----QKSRVLANQYTDIIIDERRKSLQAGV 273
QY      288 GEECKEPAK--DLIGLMIQAKNVTVQ---DIVECKSPFFAGKQTTSNLITWTITLL 340
DB      274 KOD-NTPKRYQODEFLDYLSAKDESGSSPSDIDVHSEVSTFLLAGHDTLAASISWILYCL 332
QY      341 SMHPWOKARDEVLRVCGSRDVPKDHVVVLKTLISLINESLRLVYPIVATIRAKSDV 400
DB      333 ALNPHQRCREBRVAGIIGDGSSTTMDQLGEMSTTMCIKETCRILPAVVISIRDSKYL 392
QY      401 KL-GGYKIPCGTELLIPPIAVHND-QAIVGNDVNEFPNAPPADGVPRANHPVGFIPFGL 458
DB      393 TFPDQCTLPAGITVVLSTWGLHNPAAVWKNP-KVFDPLRFSQ-ENSDQHNPAYVYPFSA 450
QY      459 GVRTTCIGONLAILQAKTLAVMIQRTFHILAPTQAHPTVLMELLYPOHGPPIFRRLT 516
DB      451 GSRNCTGGEFNMTELKVTILAILHFRVTPDPTPLTFPHNFIKPKNGMYILAKXLS 508

RESULT 5
US-08-457-274A-25
; Sequence 25, Application US/08457274A
; Patent No. 5734086
; GENERAL INFORMATION:
; APPLICANT: Scott, Jeffrey G.
; APPLICANT: Tomita, Takashi
; TITLE OF INVENTION: Cytochrome P4501pr Gene and Its Uses
; NUMBER OF SEQUENCES: 29
```

```

CORRESPONDENCE ADDRESS:
; ADDRESSEE: Nixon, Hargrave, Devans & Doyle
; STREET: P.O. Box 1051, Clinton Square
; CITY: Rochester
; STATE: New York
; COUNTRY: USA
; ZIP: 14603
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/457,274A
; FILING DATE:
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Goldman, Michael L.
; REGISTRATION NUMBER: 30,727
; REFERENCE/DOCKET NUMBER: 19603/240 (D-1519)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 716-263-1304
; TELEFAX: 716-263-1600
; INFORMATION FOR SEQ ID NO: 25:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 504 amino acids
; TYPE: amino acid
; STRADEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Rat
; STRAIN:
; DEVELOPMENTAL STAGE: Adult
; POSITION IN GENOME:
; CHROMOSOME/SEGMENT:
US-08-457-274A-25

Query Match      16.2%; Score 441.5; DB 1; Length 504;
Best Local Similarity 28.1%; Pred. No. 1.9e-37;
Matches 141; Conservative 95; Mismatches 185; Indels 81; Gaps 19;

QY      18 VILSLVVKGMSLIMWPKIEEHFSKQIRGPPYHFFIGNVKELVGMMLKASHMPFS 77
DB      13 VLLAVLV-----LLYGGTTHGLFKQGIPIGB-----KELPF- 46
QY      78 HNLIPRLSYTH-NR-----KIYGATPIVWFG--PTRLTVADPLLREIFSK---SE 125
DB      47 ---EGVLANYMGLMKFVDECHKKYGIKGLFDGQMLPALT--DTEMKNVAVKCSFV 101
QY      126 FYEKNEAHPVVKQLEGDLISLKEKWAHHRKIIISPTFHEMNIKLVPVVLKSVTDMVK 185
DB      102 FTRNRDGPV--GIMGKAVVAADBEWKRYRALLSPFSGRLKEMFPII-EGYGDILVK 158
QY      186 WSDKLSNGEVEVDYEMFQILTEDVISRTAFG---SSYEDGAAVRLQAQOMLCAEAF 242
DB      159 YLKQEAETGK-FVTMKVFGAYSMVDYTSISFGVNDLSLNPDPF-VEKTKLLEFDF 216
QY      243 QKVFIPIGYRF--PPTGNTKSRKLDKIRKSLKLIERRQNAIDGEECKEPAKDL 299
DB      217 DPLFLSVLFPFLPIYEMINICFPKDSIEFFKRVYMKERLDS-----VQKRVDP 271
QY      300 LGIMIQAKN-----VYQDIVEECKSPFFAGKQTTSNLITWTITLLSNMP 344
DB      272 LQIMNNAHNSKDESHSTALSDEWETASII-----FIFAGYPTSTSTISFVLSHSLATP 326
QY      345 EMOAKARDEVLRVCGSRDVPKDHVVVLKTLISLINESLRLVYPIVATIRAKSDVKG 404
DB      327 DTQKLGELIDRALPNAPPIYDTVMEMELDVAVNETLRLYIIGNRLERVKCKDVEING 386
QY      405 YKIPCGTELLIPPIAVHNDQALWGNVDNEFPNAPPADGVPRANHPVGFIPFGLGVRTCT 464
```

Mon May 24 08:10:10 2004

us-09-992-901-2.rat

Page 4

Db 387 VEMPKGSVWIMPSYALHRDPQHW-DEPEEFPERFSK-ENKGSIDPYVILPFGNGPNCI 444  
Qy 465 GQNTALLOAKLTAVMIQRTF 486  
Db 445 GMRFLMMNKLTATYVQNFSE 466

RESULT 6  
PCT-US95-05758-25  
Sequence 25, Application PC/TUS9505758  
GENERAL INFORMATION:  
APPLICANT: Cornell Research Foundation, Inc.  
TITLE OF INVENTION: Cytochrome P4501pr Gene and Its  
TITLE OF INVENTION: Uses  
NUMBER OF SEQUENCES: 29  
CORRESPONDENCE ADDRESS:  
ADDRESSES: Nixon, Hargrave, Devans & Doyle  
STREET: P.O. Box 1051, Clinton Square  
CITY: Rochester  
STATE: New York  
COUNTRY: USA  
ZIP: 14603  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent in Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US95/05758  
FILING DATE:  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Goldman, Michael J.  
REGISTRATION NUMBER: 30,727  
REFERENCE/DOCKET NUMBER: 19603/241 (D-1519)  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 716-263-1304  
TELEFAX: 716-263-1600  
INFORMATION FOR SEQ ID NO: 25:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 504 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
ORGANISM: Rat  
STRAIN: Unknown  
DEVELOPMENTAL STAGE: Adult  
POSITION IN GENOME:  
CHROMOSOME/SEGMENT: Unknown  
PCT-US95-05758-25

Query Match 16.2%; Score 441.5; DB 5; Length 504;  
Best Local Similarity 28.1%; Pred. No. 1.9e-37;  
Matches 141; Conservative 95; Mismatches 185; Indels 81; Gaps 19;

Qy 18 VILSVLYVKMSLMMRPKRIEHEFSKQIRGPRPHFGVGNKELVGMMLKASSHPMFS 77  
Db 13 VLLAVLV-----LLYGFGRTHGLFKKQIGPG-----KPLPF- 46  
Qy 78 HNLIPRLSFEYH-NR-----KIYGATFLVWFG--PFRLLTVADPDLIRLIFSK--SE 125  
Db 47 ---FGLVLANVYMLNMFVDECHKKYKIGLFDGQWPLPAIT--DTEMIKVILVECPFSV 101  
Qy 126 FYEKNEAHLVKQLEBDDGLISLKGEMAHHRKIISTFMENLKLIVPVVLSKVTDMWDK 185  
Db 102 FTTRRPFQPV--GIMGAVSVADDEWKKYRALSLFTSGRLKEMFPII-EGYGDILVK 158  
Qy 186 WSDKLSNGEVEVDVEMVFOILTEDVISRTAFG---SSYEDGRAVFRLOAQOMLLCAEAF 242

Db 159 YLKDAEAGK-PVTMKVFGAYSMVDVITSTSGVNVDSLNNPDPF-VEKTKLHPDF 216  
Qy 243 QKVFIPGYR---PFRGNLKSRLDKIRKSLKILERRARONADGECEKEPAKDL 299  
Db 217 DELPISVLPFLPIPIEMINICMFPKDSIEFFKFPYRNKERLDS-----YQKRVDF 271  
Qy 300 LGMLTQAN-----VTVQDIVECKSFYFKQTSVLTWTITLISMP 344  
Db 272 LQLMNNAHNDKDKSHSTALSDMEITAOISII-----FIFAGYEPTSTSLFVLSATHP 326  
Qy 345 EMOAKARDEVLRVCGSRDVPFKDHYVVLKTLMSLINESLRYPIVATIRAKSDVLAG 404  
Db 327 DQKQGEIDRALPNKAPPTDYVMEHEYLDVMTETLLPIGNRLERCKKDVING 386  
Qy 405 YKIPGTELLPIIAVHDDQIWNQDVNEFNPAFADGVPAKHVGFIPGLGYRTCI 464  
Db 387 VEMPKGSVWIMPSYALHRDPQHW-DEPEEFPERFSK-ENKGSIDPYVILPFGNGPNCI 444  
Qy 465 GQNTALLOAKLTAVMIQRTF 486  
Db 445 GMRFLMMNKLTATYVQNFSE 466

RESULT 7  
US-09-144-367-2  
Sequence 2, Application US/09144367  
Patent No. 6432639  
GENERAL INFORMATION:  
APPLICANT: Iacheter, Jay  
APPLICANT: Guido, Marco  
TITLE OF INVENTION: GENOTYPING OF HUMAN CYP3A4  
FILE REFERENCE: SEQ-12P  
CURRENT FILING DATE: US/09/144,367  
PRIOR APPLICATION NUMBER: 1998-08-31  
PRIOR FILING DATE: 1997-09-10  
NUMBER OF SEQ ID NOS: 58  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 2  
LENGTH: 503  
TYPE: PRT  
ORGANISM: H. sapiens  
US-09-144-367-2

Query Match 15.3%; Score 418; DB 4; Length 503;  
Best Local Similarity 27.3%; Pred. No. 5.6e-35;  
Matches 139; Conservative 92; Mismatches 185; Indels 94; Gaps 18;

Qy 14 LVLSVLYSVYKMSLMMRPKRIEHE-FSKQIGRPRPHFGVGNKELVGMMLKASS 71  
Db 13 LLLAVSVLLLYG-----THSGLEFKKQIGPGT----- 42  
Qy 72 HMPFSNHLIPRLSFEYH-----HWKTYGATFLVWFGPFRLLTVADPDLIRLIFSK 123  
Db 43 -PLPFLGNILS-----YKGFQCFPMDECHKKYGVWGFYDQGVLAITPDMIKIVLVK 96  
Qy 124 ---SEFYENEAHLVKQLEBDDGLIS-----LKEKNAHHRKIISTFMENLKLIVPVV 175  
Db 97 ECVSVFTNRPPGPV-----GFWKSAISIAEBEWRRLSLSPFTSGRLKEMFPII 149  
Qy 176 LKSVTDMWDKMSDKLSNGEVEVDVEMVFOILTEDVISRTAFG---SSYEDGRAVFRLOA 232  
Db 150 AQ-YGVIVLVNRRRAEAGTK-PVTLKDVFGAYSMVDVITSTSGVNVDSLNNPDPF-VEN 206  
Qy 233 QOMLLCAEAFQKVFIPGYR---PFRGNLKSRLDKIRKSLKILERRARONADGECEKE 289  
Db 207 TKKLIRFDPDLDPFLSITVFPFLPIILVNI CVFPRVNTFLKSVRMKESRL----- 261  
Qy 290 ECKEPAKDLGLMTQAN-----VTVQDIVECKSFYFKQTSVLTWTITLISMP 340  
Db 262 EDTQKHVDFLQMLDSQNSKATESHKLSDLELVLAQSIITIFAGYETTSVLSFIMVEL 321

```

QY 341 SMEPEWAKARDEVLKVCGRDVPVTKDHWKTKLSTLINESLALPYEIVATIRAKSDV 400
DB 322 AHPEDVQKQKQSEIDALPDKAPPTVTVLQWETLMDVNETLPLFIAHRLERVCQKV 381
QY 401 KLGKYKIPCGTELLPIIAVHDDAIGNDVNEFNPAFA---DGVPRAKHPVGFIFF 456
DB 382 EINGNEIFPKGVMVMTIPSAVLRDPKYM-TSEPEKFLPERFSKXKNDI----DPIYITPF 435
QY 457 GLGVFTICIGONLAILIQAULTLAVMOTRETF 486
DB 436 GSGPRCIGMRFALMMKALIRVQNFSE 465

RESULT 8
US-08-948-564-16
; Sequence 16, Application US/08948564
; Patent No. 6121512
; GENERAL INFORMATION:
; APPLICANT: Simmszy, Balazs
; APPLICANT: Dewey, Ralph E.
; APPLICANT: Cordin, Frederick T.
; TITLE OF INVENTION: No. 6121512: Cytochrome P-450 Constructs and
; TITLE OF INVENTION: Methods of Producing Herbicide-Resistant Transgenic Plants
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Virginia C. Bennett
; STREET: PO Box 37428
; CITY: Raleigh
; STATE: No. 6121512th Carolina
; COUNTRY: USA
; ZIP: 27627
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/948,564
; FILING DATE:
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Bennett, Virginia C.
; REGISTRATION NUMBER: 37,092
; REFERENCE/DOCKET NUMBER: 5051-409
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919-854-1401
; TELEFAX: 919-854-1401
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 576 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: Protein
; US-08-948-564-16

Query Match 14.8%; Score 404; DB 3; Length 576;
Best Local Similarity 26.8%; Pred. No. 2.1e-33;
Matches 133; Mismatches 93; Indels 201; Gaps 19;

```

```

DB 238 GSATKESPVIAKAY-----GTLPEAHESTYFIPIYWKILARWIVPROKFDQDLKINT 292
QY 271 SLKLTERRRQNAIDSGECKER---AAKD--LIGLMIQAKNTVQD--IVECKSGFF 323
DB 293 CLDGLRNNAESRQETDVEKQLQORDYLNIKDSALRFLVDRGADVDRQRLDMLMLI 352
QY 324 AGKOTSNLITWTITLLSMEPEWAKARDEVLKVCGRDVPVTKDHWKTKLSTLINESL 383
DB 353 AGHETIAAVTWAVFLAQNPSKMKKAQAVDVLGT--GRPTFESKELQYIRLIVREAL 411
QY 384 RLIPPIVATIRRA-KSDVYKGGYK-----IPCTELLPIIAVHDDAIGNDVNEFN 435
DB 412 RLVPQPLRLRLSKSDVLPQGHKQKQGYALPAQTDVFIQVNLHRSPPYFMDRP--DDFE 470
QY 436 PARF-----ADG-----VPRAKHPVGFIFFGLGVFTICIGONLAILIQA 474
DB 471 PERFLVQNKHEIEGAGLDPBSRFGALYFNEVISDFALFFGGGRKCVGDQFALWEST 530
QY 475 LTLAVMIQRTFHLAFT 491
DB 531 VALTMLQWFDVELKGT 547

RESULT 9
US-09-583-447A-2
; Sequence 2, Application US/09583447A
; Patent No. 6645745
; GENERAL INFORMATION:
; APPLICANT: WOJNOMSKL, Leszek
; APPLICANT: GELINER, Klaus
; APPLICANT: EISELT, Regina
; TITLE OF INVENTION: IDENTIFICATION OF A NEW MEMBER OF THE CYTOCHROME P450 3A
; TITLE OF INVENTION: (CYP3A) GENE FAMILY: CYP3A
; FILE REFERENCE: 310115.401
; CURRENT APPLICATION NUMBER: US/09/583,447A
; CURRENT FILING DATE: 2000-05-30
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 2
; LENGTH: 503
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-583-447A-2

Query Match 14.1%; Score 386; DB 4; Length 503;
Best Local Similarity 25.9%; Pred. No. 1.3e-31;
Matches 128; Mismatches 93; Indels 202; Gaps 72; Gaps 16;

```

```

Db      336 DAVLEPNKAPVTDALVQMEYLDVYNETLRLFPVSVTRVCKDKDLEINVFIPKGLAVM 395
Qy      415 IPIIVHDDQIMGNDVNEFPNAPADGVPRAAHPVG---FIPFGIGVTCIGOMLAI 471
Db      396 VPIYALHDDPKYM-TEPEKFCPEBRS---KKNDSDIDLYRIPIFGAGPANCIGMFAL 450
Qy      472 QAKTLAVMIQRTF 486
Db      451 NIKLAVIRALQNFSS 465

```

```

RESULT 10
US-09-583-447A-4
; Sequence 4, Application US/09583447A
; Patent No. 6645745
; GENERAL INFORMATION:
; APPLICANT: NOJONOSKI, Ieszek
; APPLICANT: GELLMER, Klaus
; APPLICANT: EISELT, Regina
; TITLE OF INVENTION: IDENTIFICATION OF A NEW MEMBER OF THE CYTOCHROME P450 3A
; FILE REFERENCE: 310.115.401
; CURRENT APPLICATION NUMBER: US/09/583,447A
; CURRENT FILING DATE: 2000-05-30
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 504
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-583-447A-4

```

```

Query Match      13.8%; Score 377.5; DB 4; Length 504;
Best Local Similarity 25.8%; Pred. No. 1e-30;
Matches 128; Conservative 94; Mismatches 20; Indels 73; Gaps 17;

Qy      22 LVIVKMSLLMWRPRKIEEHFSKQIRGPPYHFIQNVLELVGMMLKASHPMPESHNL 81
Db      13 VLVATSLVLYITGTSKILFKLGIPGT-----PLPILGTL 51
Qy      82 PVLSEFYHWRKI---YGATFLWFGPTFLVADPDILREIFSK---SEFEKNEANPL 135
Db      52 FYLRGLMNFRECKEKGEMWGLYEGQCPMLVIMPDMIKTVLVECVSVFNQPLGPM 111
Qy      136 VKOLEGGGLS-----LKGEKMAHHRKIISPFHEMLKLVAVVTKSVTDVDEKMDKL 190
Db      112 -----GFLKSAISPADEDEMKRITLSPAFTSVKFKEMVPII-SQCGMLVNSLROE 163
Qy      191 SENGVEVVDYVEFOILTEVDVISRIFG---SSYEDGRAVFLQAOQMLCAEAFQYFI 247
Db      164 AENSK-SINKDFGAYTMDVITGLFGVLDISLNNPDPF-LKMMKLLKLD-FLDPFL 220
Qy      248 PGYRFF---PTRGNLKRKLDEIKRSILKLIERRROVAIDGEGCEKPAKOLLGIM 303
Db      221 LLISLFPFLTFVEALNIGLFPDOVTHFLNLSIERKESRL-----KOKQRVDVFPQW 275
Qy      304 IOAKN-----VTQDIVECKSEFFAGKQTSNLLTWTTILLSMPEWQAARDEV 354
Db      276 IDGNSKETSKHKLSDLELVAQSIIIFAAVDTSTLPLFIYELATHPDVQOKLOEII 335
Qy      355 LRVCGSRDVPFKDHYVTKLTLMSILNESLALYPIYATIRAKSDVKGLGYKIPGTELL 414
Db      336 DAVLEPNKAPVTDALVQMEYLDVYNETLRLFPVSVTRVCKDKDLEINVFIPKGLAVM 395
Qy      415 IPIIVHDDQIMGNDVNEFPNAPADGVPRAAHPVG---FIPFGIGVTCIGOMLAI 470
Db      396 VPIYALHDDPKYM-TEPEKFCPEBRS---KKNDSDIDLYRIPIFGAGPANCIGMFAL 450
Qy      471 LQAKTLAVMIQRTF 486
Db      451 NIKLAVIRALQNFSS 465

```

```

RESULT 11
US-09-518-386B-1
; Sequence 1, Application US/09518386B
; Patent No. 6365386
; GENERAL INFORMATION:
; APPLICANT: HOSHINO, Tatsuo
; APPLICANT: OJIMA, Kazuyuki
; APPLICANT: SETOGUCHI, Yutaka
; TITLE OF INVENTION: ASTAXANTHIN SYNTHETASE
; FILE REFERENCE: ASTAXANTHIN SYNTHETASE
; CURRENT APPLICATION NUMBER: US/09/518,386B
; CURRENT FILING DATE: 2000-03-03
; PRIOR APPLICATION NUMBER: EP 99104668.1
; PRIOR FILING DATE: 1999-03-09
; PRIOR APPLICATION NUMBER: EP 00101666.6
; PRIOR FILING DATE: 2000-02-01
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 557
; TYPE: PRT
; ORGANISM: Phaffia rhodozyma
; FEATURE:
; NAME/KEY: TRANSIT
; LOCATION: (1)..(26)
US-09-518-386B-1

```

```

Query Match      13.4%; Score 367; DB 4; Length 557;
Best Local Similarity 22.2%; Pred. No. 1.5e-29;
Matches 128; Conservative 111; Mismatches 203; Indels 134; Gaps 21;

Qy      13 VLVSVILSVYKMSL---LWMPRKIEEHFSKQIRGPPY-HFIQNVLELVGMML 67
Db      4 LVILGALGLAFAFWSAIFSLYLAAPR---SLVNLQCPNHTNFTFGFDILS--- 56
Qy      68 KASHPMPESHNLRLVLSFYHWRKIYGATFL---WFGPTFLVADPDILREIFSKS 124
Db      57 -----ARTGEHAKYREKYGSLTRAGTAGAVLNST--DPVFNHWKEA 100
Qy      125 EPEYK-NEAHPVVKOLEGGGLSLKGEKMAHHRKIISPFHEMLKLVAVVTKSVTDMV 183
Db      101 YDIPKGMARALRLATGQVYTAGBAHKHRRIMISLSAQAVKSNVPFIFLEKMELV 160
Qy      184 DKMSDKLSN-----GEVE-----VDYVEWFOILTEVDVISRIFGSSYE----- 222
Db      161 DKMDEDAEKMDAVGESAGEKKATRLTEGVADVCKDVGRAFLDMALAGFYKSDSLQNK 220
Qy      223 -----DGRAVFLQAOQMLCAEAFQYV---FIPGYRFF-----TRGNLK 260
Db      221 TNELYAFGLDGRAP-----TLDSFKAIMDFVYFRTMKRRHEIPLTQGLAV 270
Qy      261 SRKLDKIRKSLKLIERRROVAIDGEGEC---KEPAKOLLGIMQIA-----KN 308
Db      271 SRRVG-----IELMEQKQAVLSASDQAVDKKQOVGGRDLSLVANANALPESQK 323
Qy      309 VTYQDIVECKSEFFAGKQTSNLLTWTTILLSMPEWQAARDEVLRVCGSRDVPFKD 368
Db      324 LSDEEVLAQISNLPRGYTSSVLTLMFHRISSEKAVAQDKLREICQI--DTMPFLIDE 381
Qy      369 VTKLTLMSILNESLALYPIYATIRAKSDVK-----LGQYKIPCGTELL 413
Db      382 LNLPLYLEAFVYESLELDPSPYANRECKEDFIPLAEPYIGDGSVINEVRLITKTMV 441
Qy      414 IPIIVHDDQIMGNDVNEFPNAPADGVPRAAK-----HPVGIPIFGIGVTCIGQ 466
Db      442 MLPFNINSKFIYGDABEFPERKLEBVTDSINSIAPYGHQSFIS---GPRACGM 498
Qy      467 NLAILOAKTLAVMIQRTF--LAFTYQHAFTVL 499
Db      499 RFAVAEWKAFLEVTLRVQFEPLISHPEYEHLLII 534

```

Mon May 24 08:10:10 2004

us-09-992-901-2.ral

Page 7

RESULT 12  
US-09-518-386B-3  
; Sequence 3, Application US/09518386B  
; Patent No. 636386  
; GENERAL INFORMATION:  
; APPLICANT: HOSHINO, Tatsuo  
; APPLICANT: OIWA, Kazuyuki  
; APPLICANT: SETOGUCHI, Yutaka  
; TITLE OF INVENTION: ASTAXANTHIN SYNTHETASE  
; FILE REFERENCE: ASTAXANTHIN SYNTHETASE  
; CURRENT APPLICATION NUMBER: US/09/518,386B  
; PRIOR FILING DATE: 2000-03-03  
; PRIOR FILING DATE: 1999-03-09  
; PRIOR APPLICATION NUMBER: EP 00101666.6  
; PRIOR FILING DATE: 2000-02-01  
; NUMBER OF SEQ ID NOS: 32  
; SOFTWARE: Patentln Ver. 2.1  
; SEQ ID NO 3  
; LENGTH: 557  
; TYPE: PRT  
; ORGANISM: Phaffia rhodozyma  
US-09-518-386B-3

Query Match 13.4%; Score 367; DB 4; Length 557;  
Best Local Similarity 22.2%; Pred. No. 1.5e-29;  
Matches 128; Conservative 111; Mismatches 203; Indels 134; Gaps 21;

QY 13 VLVSLVSLVIVKMSL---LWRRPKIEHRSKQIRGPRY-HFETGVKELVGMVL 67  
DB 4 LVLLTALGALPASMALFSLVLAERR---SLVNLQSPNTNFTSNFIDLS--- 56  
QY 68 KASHMPSPSHNLPVLSFYHMRKIKGATFL--VWGFPTLVADBDLREIFSKS 124  
DB 57 -----ARTGEHAKYREKXGSLFRAGAGAVLNST--DKVFNHMKKA 100  
QY 125 EPEYK-NEAPLVKQLEGGLSLGGERMAHHRKISTFEMELKLVVVLKSTIDWV 183  
DB 101 YDPKPMARVLRATGCVVTAEGEAKHRRKIMTSLAQVKSMMVIFLEKMEVL 160  
QY 184 DKMSDKLSEN-----GEVY-----VDVYEWFOILTEDVISRTAFSSYE----- 222  
DB 161 DKMEDPAEKDVAAGSAGEKATLETGVVDKDWGRATLDVWALAGDYKSDGLQNK 220  
QY 223 -----DGAARVLOAQVLLCAEFQKY--FIRGYRFP-----TGKNIK 260  
DB 221 TNELVYAVGLTDGPRP-----TLDSFKAIMDFVYFRTMKRHEIPLTQGLAV 270  
QY 261 SRKLDKIRKSLKLIERRQNAIDGEGEC--KEPAKDLGLMIQA-----KN 308  
DB 271 SRVVG-----TELMEQKQAVLGSASDAVDKDDVQGRDILSLVAVANIANLPESQK 323  
QY 309 VVVQDIVECKSPFFPGKQTTNLTWTITLLSWHPWQAKARDEVLRVCGSDVPTKH 368  
DB 324 LSEBEVLAQISNLRFGYETSSVLTMMFRLSDEKRAVDKLRBEICQI--DTDMPLD 381  
QY 369 VVRLKTLISMLNESRLVPPVATIRAKSDVK-----LGGYKIPQGTETL 413  
DB 382 LNLPLFLAFAVKESRLDPPSPYANREGLKDEDFILAPVYIGDGSVINEVITGTWV 441  
QY 414 LPIIIVHNDQAIWGDVNEFPNAPFADGVPRAK-----HPVGGIPRGLGVRCTIQ 466  
DB 442 MLEPLNINSKFTYGDSEEFREPERMLDVTDSLANSIAPYGHQAGFIS--GPRACFG 498  
QY 467 NLAIOAKLTLAVMIOQFTFH--LAPTYOAPTVL 499  
DB 499 RFAVEMKAFLEVTLRVQFEBIISHPEYEHITLI 534

RESULT 13  
US-09-904-615-69  
; Sequence 69, Application US/09904615  
; Patent No. 6566325

; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: 49 Human Secreted Proteins  
; FILE REFERENCE: P2032P1  
; CURRENT APPLICATION NUMBER: US/09/904,615  
; CURRENT FILING DATE: 2001-07-16  
; PRIOR APPLICATION NUMBER: 09/511,554  
; PRIOR FILING DATE: 2000-02-23  
; PRIOR APPLICATION NUMBER: 60/097,917  
; PRIOR FILING DATE: 1998-08-25  
; PRIOR APPLICATION NUMBER: 60/098,634  
; PRIOR FILING DATE: 1998-08-31  
; NUMBER OF SEQ ID NOS: 170  
; SOFTWARE: Patentln Ver. 2.0  
; SEQ ID NO 69  
; LENGTH: 502  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: SITE  
; LOCATION: (502)  
; OTHER INFORMATION: Xaa equals stop translation  
US-09-904-615-69

Query Match 12.7%; Score 347; DB 4; Length 502;  
Best Local Similarity 26.0%; Pred. No. 1.6e-27;  
Matches 129; Conservative 87; Mismatches 196; Indels 84; Gaps 22;

QY 49 GPRHPFPGVKEVGMMLKASHMPSPSHNLPVLSFYHMRKIKGATFLVWGFPT 108  
DB 42 GPRGLPFGNIVSLA-----ASSP-----LPHV--YMKQSVYEISLGLGIST 86  
QY 109 LTVADPDLIRE-IFSKEPEYKNEAHP-LVKQEGDGLSLK-GEKMAHHRKISPT 164  
DB 87 VLVNGDVVVECLVHQSIFADRPCLEFMKVTXKGLNSRYGSGVDRRLAVNSERY 146  
QY 165 -----MENLKLVVVLKSTIDWVDKMSDKLSENGEVEDVYEWFOILTEDVISRT--AF 217  
DB 147 FGVQGSFBSKILIEFKFNDAIETYGR-----PDRFKLIINAVSNITNLIIF 196  
QY 218 GS--SYEDG--RAVFLQAOQVLLCAEFQKYFIPGYRFPFTRGKSRKLDKIRKS-- 271  
DB 197 GERFTEYEDTFQMIELFSENVELAASA--SVFL--YNAFPWIGILPFGKHQQLFRNAAV 252  
QY 272 -----LKLTERR-----QNAIDGEGEKEPAADLGLMIOAKN-----VTVODI 314  
DB 253 VYDFLSRLKASVNEKPOLQHFVDAYLD-----NDQKNDSSSTFSKENTL 300  
QY 315 VEECKSPFGKQTTNLTWTITLLSWHPWQAKARDEVLRVCGSDVPTKHVVKLKT 374  
DB 301 IFSVGEILTAGTETTNVLRMAILFMALYPNIGQVQJEDILMGPNKSPWDXCKMPY 360  
QY 375 LSMILNESRLVPPVATIRAKS--DVYLGKVPICGHELLPIIIVHNDQAIWGDVNE 433  
DB 361 TEAVLHVLAFQCNIVLGIFFAITSDDAVRGISIKGTIVTINLVSHPFDKXN-RDPEV 419  
QY 434 FNPAPADGVPRAKHVGFIIPFGAGVTCIGQNTAILIOAKLTLAVMIOQFTFH--LA 489  
DB 420 FHEPERFDSGGYPAKKE-ALVPSLGRHCHGEHARWEMPLFTALORFHLHPHELV 478  
QY 490 PTYOHAPTVMLLTPQ 505  
DB 479 PDKR--PRIGMTLPQ 492

RESULT 14  
US-08-457-274A-23  
; Sequence 23, Application US/08457274A  
; Patent No. 5734086  
; GENERAL INFORMATION:  
; APPLICANT: Scott, Jeffrey G.  
; APPLICANT: Tomita, Takashi  
; TITLE OF INVENTION: Cytochrome P4501pr Gene and Its Uses

Mon May 24 08:10:10 2004

us-09-992-901-2.ra1

Page 8

```

NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:
ADDRESSEE: Nixon, Hargrave, Devans & Doyle
STREET: P.O. Box 1051, Clinton Square
CITY: Rochester
STATE: New York
COUNTRY: USA
ZIP: 14603

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/457,274A
FILING DATE:
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: Goldman, Michael L.
REGISTRATION NUMBER: 30,727
REFERENCE/DOCKET NUMBER: 19603/240 (D-1519)
TELECOMMUNICATION INFORMATION:
TELEPHONE: 716-263-1304
TELEFAX: 716-263-1600
INFORMATION FOR SEQ ID NO: 23:
SEQUENCE CHARACTERISTICS:
LENGTH: 507 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: drosophila melanogaster
STRAIN:
DEVELOPMENTAL STAGE: Adult
POSITION IN GENOME:
CHROMOSOME/SEGMENT:
US-08-457-274A-23
Query Match 12.6%; Score 344.5; DB 1; Length 507;
Best Local Similarity 23.2%; Pred. No. 3.1e-27;
Matches 122; Conservative 116; Mismatches 211; Indels 77; Gaps 19;
```

```

19 ILSLVIVKMSILMWRPKIEHFSKQIGRPYHFFIGNYKELVGMMLKASHPMPFSH 78
3 VILYLIALISLLAYLYHNFNVMNRGLPHDAPRLVGN--WVGRKRNVMDFY-- 57
79 NILPRVLSFYHHRKIYGATPL-WVFGPTFRITVADPLDILREIFSK--SEFEKNEAH-- 133
58 -----DYNNKRR-SGFPFVGFFLHKPAPAFIVDTOLAKNIIKDFSNFADRGQFHNG 109
134 ---PLVKQLEGGLSLGKSKVAHHRKIISTFHEHNTKLLVPVVLKSVTDMVCKSDKL 190
110 RDDLPLQH-----LFLNDGKKWKDMQRSLTFTSGKMKFMPFVIAKVSSEFVAVITQOV 164
191 --SENGEVEVDVYEMFOILTEVDVISTAFSGSYEDGR---AIFRLQAOMLCAAPQKV 245
165 PAQONQAV-LEIEMARFTTDVIGTCRFGIECNTLRTPVSDFRIMQOKVFTDRHGKLL 223
246 FIDYRFPPTRGKLDKEIRKSLKLIIE---RRQNAIDGEGECKEPPAADDL 300
224 TMVFSPKPLASRLRBMWPEVDHGFPMRLVNDTIALRERENF-----KRNDFM 272
301 GLMTQAK-----NWTYQDIVECKSFFPFGKQTSNMLTWTITLLSMRE 345
273 NLLIEIKOKGSFTLNDGEVIEGMDIGELAAQVFFVYVAGFTSSSTSYCLYELAQOD 332
346 WQAKARDEVLRVCGSRDVPDKHVVAKLT-LSMLINESLRLLPPIVATIRRAKSDVKLG 404
333 IQRLRNEIGTVLEBEGQLTYSIRAKMTYINQVISETLRLTLVPHLERKALNDYVVP 392
```

```

405 YK---IPCGTELLIPIIAVHHDDQALMGNDVNEFNPAFADGVR--AAKIPVGFIFGG 459
393 HEKVIENGIVITIPACAYHRDEDLVNP-ETDPERFS--PEKVAARSVEMLPFGDG 448
460 VRTICGNLAILQAKLTILAVMIQRFTHLAPT-----YQHAFTYML 501
449 PRNCIGRRFGOMCARIGLACIISRFRVSDTTEIPLKSPMSIVL 494
```

```

RESULT 15
PCT-US95-05758-23
Sequence 23, Application PC/TUS9505758
GENERAL INFORMATION:
APPLICANT: Cornell Research Foundation, Inc.
TITLE OF INVENTION: Cytochrome P4501pr Gene and Its
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:
ADDRESSEE: Nixon, Hargrave, Devans & Doyle
STREET: P.O. Box 1051, Clinton Square
CITY: Rochester
STATE: New York
COUNTRY: USA
ZIP: 14603

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/05758
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Goldman, Michael L.
REGISTRATION NUMBER: 30,727
REFERENCE/DOCKET NUMBER: 19603/241 (D-1519)
TELECOMMUNICATION INFORMATION:
TELEPHONE: 716-263-1304
TELEFAX: 716-263-1600
INFORMATION FOR SEQ ID NO: 23:
SEQUENCE CHARACTERISTICS:
LENGTH: 507 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: drosophila melanogaster
STRAIN:
DEVELOPMENTAL STAGE: Adult
POSITION IN GENOME:
CHROMOSOME/SEGMENT:
PCT-US95-05758-23
```

```

Query Match 12.6%; Score 344.5; DB 5; Length 507;
Best Local Similarity 23.2%; Pred. No. 3.1e-27;
Matches 122; Conservative 116; Mismatches 211; Indels 77; Gaps 19;
19 ILSLVIVKMSILMWRPKIEHFSKQIGRPYHFFIGNYKELVGMMLKASHPMPFSH 78
3 VILYLIALISLLAYLYHNFNVMNRGLPHDAPRLVGN--WVGRKRNVMDFY-- 57
79 NILPRVLSFYHHRKIYGATPL-WVFGPTFRITVADPLDILREIFSK--SEFEKNEAH-- 133
58 -----DYNNKRR-SGFPFVGFFLHKPAPAFIVDTOLAKNIIKDFSNFADRGQFHNG 109
134 ---PLVKQLEGGLSLGKSKVAHHRKIISTFHEHNTKLLVPVVLKSVTDMVCKSDKL 190
110 RDDLPLQH-----LFLNDGKKWKDMQRSLTFTSGKMKFMPFVIAKVSSEFVAVITQOV 164
```



Mon May 24 08:10:10 2004

us-09-992-901-2.rtf

Page 9

QY 191 --SENBEVDVTEMFQIITEDVISTRAFSSYEDGR---AVFRLQAOQMLCAEAFQKV 245  
Db 165 PAQNGAV--LEIKELMARFTTDVIGTCRFGIECNLTLPVSDFRMTGQKVFTDMRGKLL 223  
QY 246 FIDGYRFPPTRGKLSKRLDKXIRKSLKLE----RRRONAIDGEGECKEPAXKDL 300  
Db 224 TWFVPSFPKLASRLRMRMPEDVHQPFMLVNDTLALRENF-----KXNDPM 272  
QY 301 GLMIQAK-----NVTQDIVEECKSFFFAQQTNSLLTWTTLISMPHE 345  
Db 273 NLLIELKQKGSFTLNDGEVIEGMDIGELAQVFWYVAGFETSSSTWSYCLYELAQND 332  
QY 346 WQAKARDEVLRVCGSRDVPYKDHVYKLT--LSMILNESRLYEPYATIRAKSDPKLG 404  
Db 333 IQDRLNVEIQTVLEBQGGQLYTESIKAMTYLNQVISETLKLYTLVPHLERKALNDYVPG 392  
QY 405 YK--IPCGETELLPIIAVHDOAINQNDVNEENPARFADGVPR--AAKHVYGFIPGIG 459  
Db 393 HEKLVIEKGTQVILIPACAYRDEDLXNP-ETFDPERFS--PEKVAABESVEMLPFGDG 448  
QY 460 VRTICIGONTALIQAKTLTAVMIQFTFHLAPT---YQHAPVYML 501  
Db 449 PRNCIGRRFGOMQARIGLAQIISRFRVSVDTEIPLKXSPMSIVL 494

Search completed: May 19, 2004, 10:28:27  
Job time : 24 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 19, 2004, 10:27:26 ; Search time 49 Seconds  
(without alignments)

2952.979 Million cell updates/sec

Title: US-09-992-901-2  
Perfect score: 2730

Sequence: 1 MEESSSWFIPKVLVSVIL.....LYPGGATITFRRLTHED 520

Scoring table:

BIOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1145568 seqs, 278261457 residues

Total number of hits satisfying chosen parameters: 1145568

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing:

Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

```
1: /cgn2_6/ptodaca/2/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodaca/2/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodaca/2/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodaca/2/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodaca/2/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodaca/2/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodaca/2/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodaca/2/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodaca/2/pubpaa/US09_PUBCOMB.pep:*
10: /cgn2_6/ptodaca/2/pubpaa/US09C_PUBCOMB.pep:*
11: /cgn2_6/ptodaca/2/pubpaa/US09C_NEW_PUB.pep:*
12: /cgn2_6/ptodaca/2/pubpaa/US09C_PUBCOMB.pep:*
13: /cgn2_6/ptodaca/2/pubpaa/US10_PUBCOMB.pep:*
14: /cgn2_6/ptodaca/2/pubpaa/US10C_PUBCOMB.pep:*
15: /cgn2_6/ptodaca/2/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodaca/2/pubpaa/US10_NEW_PUB.pep:*
17: /cgn2_6/ptodaca/2/pubpaa/US60_NEW_PUB.pep:*
18: /cgn2_6/ptodaca/2/pubpaa/US60_PUBCOMB.pep:*
```

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2730	100.0	520	US-09-992-901-2	Sequence 2, Appl1
2	1386.5	50.8	533	US-10-424-599-202258	Sequence 202258,
3	1165.5	42.7	547	US-10-425-114-61218	Sequence 61218, A
4	1165.5	42.7	560	US-10-425-114-59349	Sequence 59349, A
5	1165.5	42.7	562	US-10-425-114-59350	Sequence 59350, A
6	1155.5	42.3	547	US-10-425-114-65616	Sequence 65616, A
7	1138.5	41.7	524	US-10-424-599-183328	Sequence 183328,
8	1124	41.2	508	US-10-424-599-195559	Sequence 195559,
9	1114	40.8	533	US-10-424-599-269344	Sequence 269344,
10	1103.5	40.4	537	US-10-425-114-38180	Sequence 38180, A
11	1102	37.5	524	US-10-424-599-217648	Sequence 217648,
12	1024	30.5	520	US-10-424-599-228819	Sequence 228819,
13	989	26.2	458	US-10-425-114-50977	Sequence 50977, A
14	973.5	25.7	532	US-10-425-114-58587	Sequence 58587, A
15	967	25.4	540	US-10-425-114-61220	Sequence 61220, A

16	967	35.4	552	US-10-425-114-64681	Sequence 64681, A
17	962	35.2	541	US-10-425-114-61219	Sequence 61219, A
18	930.5	34.1	527	US-10-425-114-67055	Sequence 67055, A
19	894	32.7	455	US-10-424-599-247367	Sequence 247367,
20	858	31.4	388	US-10-424-599-197466	Sequence 197466,
21	664.5	24.3	298	US-10-425-114-48579	Sequence 48579, A
22	660.5	24.2	546	US-10-424-599-219019	Sequence 219019,
23	543.5	19.9	505	US-09-910-689-307	Sequence 307, App
24	543.5	19.9	505	US-10-058-270A-2	Sequence 2, Appl1
25	543.5	19.9	505	US-10-296-606-10	Sequence 10, Appl
26	541	19.8	505	US-10-010-742-307	Sequence 307, App
27	541	19.8	286	US-10-424-599-280555	Sequence 280555,
28	540.5	19.8	505	US-10-219-535-30	Sequence 30, Appl
29	540.5	19.8	505	US-10-232-230-30	Sequence 30, Appl
30	540.5	19.8	505	US-10-233-224-30	Sequence 30, Appl
31	540.5	19.8	505	US-10-067-6668-2	Sequence 2, Appl1
32	540.5	19.8	505	US-10-227-884-30	Sequence 30, Appl
33	540.5	19.8	505	US-10-230-163-30	Sequence 30, Appl
34	540.5	19.8	505	US-10-230-338-30	Sequence 30, Appl
35	540.5	19.8	505	US-10-218-631-30	Sequence 30, Appl
36	540.5	19.8	505	US-10-230-414-30	Sequence 30, Appl
37	540.5	19.8	505	US-10-216-159A-30	Sequence 30, Appl
38	540.5	19.8	505	US-10-218-849-30	Sequence 30, Appl
39	540.5	19.8	505	US-10-227-873-30	Sequence 30, Appl
40	540.5	19.8	505	US-10-227-883-30	Sequence 30, Appl
41	540.5	19.8	505	US-10-219-076-30	Sequence 30, Appl
42	540.5	19.8	505	US-10-230-434-30	Sequence 30, Appl
43	540.5	19.8	505	US-10-219-003-30	Sequence 30, Appl
44	540.5	19.8	505	US-10-219-075-30	Sequence 30, Appl
45	540.5	19.8	505	US-10-219-464-30	Sequence 30, Appl

## ALIGNMENTS

```
RESULT 1
US-09-992-901-2
; Sequence 2, Application US/09992901
; Patent No. US20020073446A1
; GENERAL INFORMATION:
; APPLICANT: Neff, Michael M.
; TITLE OF INVENTION: GENETICALLY MODIFIED PLANTS HAVING
; FILE REFERENCE: SAKIKINS, 024DVI
; CURRENT FILING DATE: 2001-11-14
; PRIOR APPLICATION NUMBER: US 09/527, 073
; PRIOR FILING DATE: 2000-03-16
; PRIOR APPLICATION NUMBER: US 60/124570
; PRIOR FILING DATE: 1999-03-16
; PRIOR APPLICATION NUMBER: US 60/170, 931
; PRIOR FILING DATE: 1999-12-14
; PRIOR APPLICATION NUMBER: US 60/172, 832
; PRIOR FILING DATE: 1999-12-20
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 520
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-09-992-901-2
Query Match 100.0%; Score 2730; DB 9; Length 520;
Best local similarity 100.0%; Pred. No. 5e-266;
Matches 520; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MEESSSWFIPKVLVSVILVYVGMGLMRRPKTIEHFQKQIRGPPHFFGNK 60
DB 1 MEESSSWFIPKVLVSVILVYVGMGLMRRPKTIEHFQKQIRGPPHFFGNK 60
QY 61 ELYGMMKASHMPFESHNLIPVLSFYHMRIKYATFVWFGPTFRVLVADPDILREI 120
```

Db 61 ELVGMMLKASSHPMPFSNNILPRVLSFYHNRKIYGATFLVWPGPTFRLTVADPDILREI 120  
Qy 121 PFSKSEFYENKNAHPVLVKOLEBDGLSLKGEKMAHHRKIIPTFEMENLKLLVAVLKSVT 180  
Db 122 PFSKSEFYENKNAHPVLVKOLEBDGLSLKGEKMAHHRKIIPTFEMENLKLLVAVLKSVT 180  
Qy 181 DMVDKMSDKLSNGSEVEVDYVEMFQILTEDVISTPAFGSSYEDGRAVFRLOAQOMMLCAE 240  
Db 181 DMVDKMSDKLSNGSEVEVDYVEMFQILTEDVISTPAFGSSYEDGRAVFRLOAQOMMLCAE 240  
Qy 241 AFQKVFIPGYRFFPRGNLKSRLDKERKSLKLIBRRONAIDGGECKEPAKADLL 300  
Db 241 AFQKVFIPGYRFFPRGNLKSRLDKERKSLKLIBRRONAIDGGECKEPAKADLL 300  
Qy 301 GLMIQAKNTVODIVECKSPFFAGKQTTSNLITWTITLLSMHEWQAKARDEVLRVCGS 360  
Db 301 GLMIQAKNTVODIVECKSPFFAGKQTTSNLITWTITLLSMHEWQAKARDEVLRVCGS 360  
Qy 361 RVPTKRDHYVVKLTISMLINESLRYPIVATIRAKSDVKGXKIPCGTELLPIIAV 420  
Db 361 RVPTKRDHYVVKLTISMLINESLRYPIVATIRAKSDVKGXKIPCGTELLPIIAV 420  
Qy 421 HHDOAIWGNVDVEFPAPFADGVPRAKHPGFIPIGIGVATCIGOMLAILOAKLTIAV 480  
Db 421 HHDOAIWGNVDVEFPAPFADGVPRAKHPGFIPIGIGVATCIGOMLAILOAKLTIAV 480  
Qy 481 IORFTFHLPYQHPATVLMILYPOHGAPIFRRLTNHED 520  
Db 481 IORFTFHLPYQHPATVLMILYPOHGAPIFRRLTNHED 520

RESULT 2  
US-10-424-599-202258  
; Sequence 202258, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424.599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 202258  
; LENGTH: 353  
; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (1)..(353)  
; OTHER INFORMATION: unsure at all xaa locations  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MFT3847\_24664C.1.pep  
US-10-424-599-202258

Query Match 50.8%; Score 1386.5; DB 12; Length 353;  
Best Local Similarity 76.8%; Pred. No. 9.1e-131;  
Matches 214; Conservative 29; Mismatches 41; Indels 13; Gaps 5;

Qy 1 MEESSSWFIPRVLVLSILVIVKMSLLMWRPKIEHFSKQIGRPPYHPIGNVK 60  
Db 1 MEESSSWFIPRVLVLSILVIVKMSLLMWRPKIEHFSKQIGRPPYHPIGNVK 60  
Qy 61 ELVGMMLKASSHPMPFSNNILPRVLSFYHNRKIYGATFLVWPGPTFRLTVADPDILREI 120  
Db 61 ELVGMMLKASSHPMPFSNNILPRVLSFYHNRKIYGATFLVWPGPTFRLTVADPDILREI 120  
Qy 56 ELVGMMLKASSHPMPFSNNILPRVLSFYHNRKIYGATFLVWPGPTFRLTVADPDILREI 117  
Db 56 ELVGMMLKASSHPMPFSNNILPRVLSFYHNRKIYGATFLVWPGPTFRLTVADPDILREI 117  
Qy 121 F-SKSEFYENKNAHPVLVKOLEBDGLSLKGEKMAHHRKIIPTFEMENLKLLVAVLKSV 179  
Db 121 F-SKSEFYENKNAHPVLVKOLEBDGLSLKGEKMAHHRKIIPTFEMENLKLLVAVLKSV 179

Qy 180 TDMDKMSDKLSNGSEVEVDYVEMFQILTEDVISTPAFGSSYEDGRAVFRLOAQOMMLCA 239  
Db 178 VEMLEKMS-AMGEKGVIEVSEMFQSLTEDVITRTAFSSYEDGRAVFRLOAQOMMLCA 236  
Qy 240 EAFQKVFIPGYRFFPRGNLKSRLDKERKSLKLIBRRONAIDGGECKEPAKADLL 299  
Db 237 DAFQKVFIPGYRFFPRGNLKSRLDKERKSLKLIBRRONAIDGGECKEPAKADLL 295  
Qy 300 LGMLIOA-----KNTVODIVECKSPFFAGKQTTSNLITWTITLLSMHEWQAK 349  
Db 296 LGMLIOASNNMNSNTVDDMVBECKSPFFAGKQTTSNLITWTITLLSMHEWQAK 352

RESULT 3  
US-10-425-114-61218  
; Sequence 61218, Application US/10425114  
; Publication No. US20040034888A1  
; GENERAL INFORMATION:  
; APPLICANT: Liu, Jindong  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Screen, Steven E.  
; APPLICANT: Tabaska, Jack E  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(5313)B  
; CURRENT APPLICATION NUMBER: US/10/425.114  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 73128  
; SEQ ID NO 61218  
; LENGTH: 547  
; TYPE: PRT  
; ORGANISM: Oryza sativa nipponbare  
; FEATURE:  
; OTHER INFORMATION: Clone ID: LIB3431-054-B3\_FIL.pep  
US-10-425-114-61218

Query Match 42.7%; Score 1165.5; DB 12; Length 547;  
Best Local Similarity 44.1%; Pred. No. 3.5e-108;  
Matches 231; Conservative 97; Mismatches 175; Indels 21; Gaps 7;

Qy 5 SSSWFIPRVLVLSILVIVKMSLLMWRPKIEHFSKQIGRPPYHPIGNVK 63  
Db 32 ASQWILMAAAVAALLMLAVSTLEWAMTPTLRLEALAQGIRGRVRLFGVDPEV 91  
Qy 64 GMMLKASSHPMPFS-NHILPRVLSFYHNRKIYGATFLVWPGPTFRLTVADPDILREI 122  
Db 92 RLNRBARKKPLPLGCHDILPRVLPWSKAVESGKSFVPGTPVMSIDPSISIEVNS 151  
Qy 123 -KSEFYENKNAHPVLVKOLEBDGLSLKGEKMAHHRKIIPTFEMENLKLLVAVLKSVT 181  
Db 152 NKFGHYGKPKRTSLGKLL-ASGVVSYGEGMAHRRILNPAFHHEKIKMFLPVFSNCTE 210  
Qy 182 DMVDKMSDKLSNGSEVEVDYVEMFQILTEDVISTPAFGSSYEDGRAVFRLOAQOMMLCAE 241  
Db 211 MVRMENSMSIEGSEVEVDYVEMFQILTEDVISTPAFGSSYEDGRAVFRLOAQOMMLCAE 270  
Qy 242 FQKVFIPGYRFFPRGNLKSRLDKERKSLKLIBRRONAIDGGECKEPAKADLL 301  
Db 271 FRITFIPOYWLPLPKNNRRRLREIEREVSRLRGIIIGRERAING-----ETSGDILG 324  
Qy 302 LMIQAK-----NVTYQDIVECKSPFFAGKQTTSNLITWTITLLSMHEWQAK 351  
Db 302 LMIQAK-----NVTYQDIVECKSPFFAGKQTTSNLITWTITLLSMHEWQAK 351  
Qy 351 LIVESNMBESNGKALQMTDEILIECKLFYFAGMETSVLLTWITLIVSMHEWQAK 384  
Db 352 DEVRVCGSRDVPKRDHYVVKLTISMLINESLRYPIVATIRAKSDVKGXKIPCGT 411  
Qy 385 EELVHNG-RTPPYDLSRLKIVTMLVVLVLPVVLTRRTYKEMELGGIKTFAV 443  
Db 412 ELLPIIAVHHDOAIWGNVDVEFPAPFADGVPRAKHPGFIPIGIGVATCIGOMLAI 471

Db 444 TLMPLFIHHDDPWGKDAEFGNPFADGISNATKYQTSFFPFMGPRICIGONFALL 503  
Qy 472 QAKTLTAVMTQRTFTHLAPTYQNAFTVLMILYPOHGAFTFRRL 515  
Db 504 EAKVAICTIIOFPSFELSPSYIHAFPTVITLHPHOGAQTILKXI 547

## RESULT 4

US-10-425-114-59349  
; Sequence 59349, Application US/10425114  
; Publication No. US20040034888A1  
; GENERAL INFORMATION:  
; APPLICANT: Liu, Jindong  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Screen, Steven E  
; APPLICANT: Tabaska, Jack E  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(5313)B  
; CURRENT APPLICATION NUMBER: US/10/425,114  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 73128  
; SEQ ID NO 59349  
; LENGTH: 560  
; TYPE: PRT  
; ORGANISM: Oryza sativa nippondare  
; FEATURE:  
; OTHER INFORMATION: Clone ID: JC-OSLEI1B347401G01\_FLI.pep  
US-10-425-114-59349

Query Match 42.7%; Score 1165.5; DB 12; Length 560;  
Best Local Similarity 44.1%; Pred. No. 3.6e-108;  
Matches 231; Conservative 97; Mismatches 175; Indels 21; Gaps 7;

Qy 5 SSSMFIPKVLVSVLSTL-VIVYKMSLTMWRPKIEHNSKOGIRGPPHYFPIGNVKELV 63  
Db 45 ASQMTLAAAAAVALMLAVSTLEWAMTPRLEBALRAGQIRNRRRLTGDVPEV 104  
Qy 64 GMLKASHPWPS-HNILPRVLSFYHWRKIYGATFLVPGPTFRITVADPDLREIFS 122  
Db 105 RLNRKARKKPLPGCHDIIIPVLPMSKAVEHGRKSPFTWPGPTFRVIMSDDESIREVWS 164  
Qy 123 -KSEFYKRNHAHPLVQLEGGDGLSLKGEKMAHRRKIISPTFMENLKLIVPVLSVTD 181  
Db 165 NKFGHYKPKPTRLGKLL-ASGVASYEGEKMAHRRILNPAFHHEKIKMLVFPNSCTE 223  
Qy 182 MYDKMSDKLSENGEVEVDYEWFOILTEDVISRTAGSSYEDGRAVFLQAOQMLCAEA 241  
Db 224 MYRWENMSIEGMSVDVWPEQNLGDIVISKTAFGSSYEEGRRIFFQLQASABRIIOA 283  
Qy 242 FQKFTIPGYFFPTRNGLSKRKLDKXIRKSLKLIERRRONALDGECECKEPAAKDLIG 301  
Db 284 FRTIFIPGYWFLPTKNNRRRLREIEREVSLLRGIIIGKRRALIKNG-----ETSGDGLG 337  
Qy 302 LMIQAK-----NVTVODIYEECKSPFPAGKQTSNLLTWITLILSMHEPQAKAR 351  
Db 338 LIVESNNRESNGKAEIGMTDEIIECKLFYFAGMETTSVLLTWITLIVSMHEPQAKAR 397  
Qy 352 DEVLAVCGSDVPTKCHVVKLTLSMLNESLRIPPIYATIRRAKSDVKGKYIPGCT 411  
Db 398 EEVLVHNG-RTPDYDSLRLKIVTWILVEVRLVPPVFLTRRYKEMELGKIXPBAV 456  
Qy 412 ELIPIYVHNDQAIKNDVNEFNPAFADGVPRAKHVPGLPGLGYRTGIGONLAL 471  
Db 457 TLMPLFIHHDDPWGKDAEFGNPFADGISNATKYQTSFFPFMGPRICIGONFALL 516  
Qy 472 QAKTLTAVMTQRTFTHLAPTYQNAFTVLMILYPOHGAFTFRRL 515  
Db 517 EAKVAICTIIOFPSFELSPSYIHAFPTVITLHPHOGAQTILKXI 560

## RESULT 5

US-10-425-114-59350  
; Sequence 59350, Application US/10425114  
; Publication No. US20040034888A1  
; GENERAL INFORMATION:  
; APPLICANT: Liu, Jindong  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Screen, Steven E  
; APPLICANT: Tabaska, Jack E  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(5313)B  
; CURRENT APPLICATION NUMBER: US/10/425,114  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 73128  
; SEQ ID NO 59350  
; LENGTH: 562  
; TYPE: PRT  
; ORGANISM: Oryza sativa  
; FEATURE:  
; OTHER INFORMATION: Clone ID: UC-OSFLCYP143F07\_FLI.pep  
US-10-425-114-59350

Query Match 42.7%; Score 1165.5; DB 12; Length 562;  
Best Local Similarity 44.1%; Pred. No. 3.7e-108;  
Matches 231; Conservative 97; Mismatches 175; Indels 21; Gaps 7;

Qy 5 SSSMFIPKVLVSVLSTL-VIVYKMSLTMWRPKIEHNSKOGIRGPPHYFPIGNVKELV 63  
Db 47 ASQMTLAAAAAVALMLAVSTLEWAMTPRLEBALRAGQIRNRRRLTGDVPEV 106  
Qy 64 GMLKASHPWPS-HNILPRVLSFYHWRKIYGATFLVPGPTFRITVADPDLREIFS 122  
Db 107 RLNRKARKKPLPGCHDIIIPVLPMSKAVEHGRKSPFTWPGPTFRVIMSDDESIREVWS 166  
Qy 123 -KSEFYKRNHAHPLVQLEGGDGLSLKGEKMAHRRKIISPTFMENLKLIVPVLSVTD 181  
Db 167 NKFGHYKPKPTRLGKLL-ASGVASYEGEKMAHRRILNPAFHHEKIKMLVFPNSCTE 225  
Qy 182 MYDKMSDKLSENGEVEVDYEWFOILTEDVISRTAGSSYEDGRAVFLQAOQMLCAEA 241  
Db 226 MYRWENMSIEGMSVDVWPEQNLGDIVISKTAFGSSYEEGRRIFFQLQASABRIIOA 285  
Qy 242 FQKFTIPGYFFPTRNGLSKRKLDKXIRKSLKLIERRRONALDGECECKEPAAKDLIG 301  
Db 286 FRTIFIPGYWFLPTKNNRRRLREIEREVSLLRGIIIGKRRALIKNG-----ETSGDGLG 339  
Qy 302 LMIQAK-----NVTVODIYEECKSPFPAGKQTSNLLTWITLILSMHEPQAKAR 351  
Db 340 LIVESNNRESNGKAEIGMTDEIIECKLFYFAGMETTSVLLTWITLIVSMHEPQAKAR 399  
Qy 352 DEVLAVCGSDVPTKCHVVKLTLSMLNESLRIPPIYATIRRAKSDVKGKYIPGCT 411  
Db 400 EEVLVHNG-RTPDYDSLRLKIVTWILVEVRLVPPVFLTRRYKEMELGKIXPBAV 458  
Qy 412 ELIPIYVHNDQAIKNDVNEFNPAFADGVPRAKHVPGLPGLGYRTGIGONLAL 471  
Db 459 TLMPLFIHHDDPWGKDAEFGNPFADGISNATKYQTSFFPFMGPRICIGONFALL 518  
Qy 472 QAKTLTAVMTQRTFTHLAPTYQNAFTVLMILYPOHGAFTFRRL 515  
Db 519 EAKVAICTIIOFPSFELSPSYIHAFPTVITLHPHOGAQTILKXI 562

## RESULT 6

US-10-425-114-65616  
; Sequence 65616, Application US/10425114  
; Publication No. US20040034888A1  
; GENERAL INFORMATION:  
; APPLICANT: Zhou, Yihua

APPLICANT: Kovalic, David K.  
APPLICANT: Screen, Steven E.  
APPLICANT: Tabaska, Jack E.  
APPLICANT: Cao, Yongwei  
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
FILE REFERENCE: 38-21(53223)B  
CURRENT APPLICATION NUMBER: US/10/425,114  
CURRENT FILING DATE: 2003-04-28  
NUMBER OF SEQ ID NOS: 73128  
SEQ ID NO 65616  
LENGTH: 547  
TYPE: PRT  
ORGANISM: Zea mays  
FEATURE:  
OTHER INFORMATION: Clone ID: 700100628\_F11.pcp  
US-10-425-114-65616

Query Match 42.3%; Score 115.5; DB 12; Length 547;  
Best Local Similarity 42.3%; Pred. No. 3.6e-107;  
Matches 223; Conservative 104; Mismatches 181; Indels 19; Gaps 5;

1 MEESSWFIPKVLVLSLVKNSLIMPRKIEHFSKQIGRPYHFFIGNVA 60  
28 MLREASPMALAGA-VASVSLMLVMTLEMAMTPEMLDRAQIGKTRRRLFTGDR 86  
61 ELVGMMLKASSHPMPS-HNLPRLSYHMRKTYGATFLVWGPTFFRLTVADPDLIR 119  
87 ETARVNRERKXPLGCHDITPRVQPMHSTIKYGLSTWFGPTPRWIDPELVK 146  
120 IFSKSEFEKNEAHPVLQLEGDLISLGEKMAHHRKISPTFHMENTKLVPVVLKSV 179  
147 VLSNFGHGKRSRRIRRLANGLVNHDEKMAHRRILNPAFHHEKIKGMMPVFSTCC 206  
180 TMVMDKSDKLSENEVDVYEWFOITLEDVTSRTAFSGSYEDGRAVFPLOQOMLCA 239  
207 IMITRMDNSMSESSSEIDVPEFQNLGVISRTAFSGSYNGRIFELQGHLERLI 266  
240 EAFQKVFIPGYRFFTRGNLKRKDKERKSLKTIERRRQNAIDEGECEKPEAKDL 299  
267 QSVQITFTIPGYFLPTKNNRBRADVAIRKILREIIKREKQD-----KNRETNDDL 320  
300 LGLMTQAK-----NVTYQDIVECKSFFPAGKQTTNLLTWTITLSNHPQAK 349  
321 LELLLESNTRQNGNASLGLTTEVDVIECKLFYFAGKMTTSVLLTWLILVSNHPQER 380  
350 ARDEVLRVCGSRDVPKDHVVLKTLTSMILNESLRLYPIVATIRAKSDVKGKIPC 409  
381 ABEVLVSHFG-RTTDPYDLSRLKTVMTLHVLRLYPLATFLRRYKHELGIGIKYPA 439  
410 GTELLPIIAVHDOAIWGNVDNEFNAPADGVPRAPKAVGVGIPRGLVRTIGONTA 469  
440 GVELLPVIFIHDDPIWGDASEFNPERFANGISSARHQAAPFPFGGGRICQSFA 499  
470 IIOAKTLTAVMIOFTFENLAPTYQHAPTYMLLYPOHGAFTFRRLT 516  
500 LIEAKMTLCTIIQRFSELSPSYTHAPTYVITLHQHQAQIRKKLS 546

RESULT 7  
US-10-424-599-183328  
Sequence 183328, Application US/10424599  
Publication No. US20040031072A1  
GENERAL INFORMATION:  
APPLICANT: La Rosa Thomas J  
APPLICANT: Kovalic David K  
APPLICANT: Zhou Yihua  
APPLICANT: Cao Yongwei  
TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
FILE REFERENCE: 38-21(53223)B  
CURRENT APPLICATION NUMBER: US/10/424,599  
CURRENT FILING DATE: 2003-04-28

NUMBER OF SEQ ID NOS: 285684  
SEQ ID NO 183328  
LENGTH: 508  
TYPE: PRT  
ORGANISM: Glycine max  
FEATURE:  
OTHER INFORMATION: Clone ID: PAT\_MRT3847\_136559C.1.pcp  
US-10-424-599-183328

Query Match 41.7%; Score 113.5; DB 12; Length 508;  
Best Local Similarity 44.6%; Pred. No. 1.7e-105;  
Matches 231; Conservative 95; Mismatches 163; Indels 29; Gaps 7;

14 LVLSVILSVI--VKGMSLIMPRKIEHFSKQIGRPYHFFIGNVA 67  
3 LLLAAVTLVFAKPFYSNFWPWRTERHFKQIGGQYRPIFGNTSEIRRLVYAKS 62  
68 KASCHPMPSHNILPRVLSFYHMRKTYGATFLVWGPTFFRLTVADPDLIRIFSKSE-F 126  
63 EASAPPPFHHIDMGRVAFYDMSRAYGTFPLWFGSTPRLAVTPEPMIKELVANTRG 122  
127 YEKNEAHPVLQLEGDLISLGEKMAHHRKISPTFHMENTKLVPVVLKSVTDVMDK 186  
123 YKVPFPQSKLFGQGLVLEBDQNALHRIINLAFNLELVGWVPDIVASVTKGLJES 182  
187 SDKLSENEVDVYEWFOITLEDVTSRTAFSGSYEDGRAVFPLOQOMLCAEAFQKVF 246  
183 EDQSGRDEFEIDVRLHDLSDVTSRTAFSGSYEGKTIENLQSQMHLFSQAVRSY 242  
247 IPGRFFPTRNLSKRLDKERKSLKTIERRRQNAIDEGECEKPEAKDLGLMTQA 306  
243 IPGRVFTKKNQKRMLEKRETSIKLIEK-----SNTENKANTVLSLMCY 293  
307 KN-----VTYQDIVECKSFFPAGKQTTNLLTWTITLSNHPQAKDEVRVCG 359  
294 KNDAGGEKLGVEEIIDECKTIYFAGKETANLLTMAILLAHQWOSVAREVHLVIG 353  
360 SRDVPKDHVVLKTLTSMILNESLRLYPIVATIRAKSDVKGKIPCSTELLPIIA 419  
354 RNRLEPADNNDLKIVTMTINETRLTPVAMLRQSKDVMGINSINIPKTYDLALTA 413  
420 VHHDOAIWGNVDNEFNAPADGV-PPAAKHVGFIPRGLVGTTCIGONTAILQAKTLA 478  
414 VHHREIIGEDYHNFMNRPSPRVRP-----VRFPGGGRICVQONLALVEAKIATA 468  
479 VMIOFTFENLAPTYQHAPTYMLLYPOHGAFTFRRLT 516  
469 LIQSYSPVLSPNVMAPIILFVTLQPOYGAIIFRKIS 506

RESULT 8  
US-10-424-599-199559  
Sequence 199559, Application US/10424599  
Publication No. US20040031072A1  
GENERAL INFORMATION:  
APPLICANT: La Rosa Thomas J  
APPLICANT: Kovalic David K  
APPLICANT: Zhou Yihua  
APPLICANT: Cao Yongwei  
TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
FILE REFERENCE: 38-21(53223)B  
CURRENT APPLICATION NUMBER: US/10/424,599  
CURRENT FILING DATE: 2003-04-28  
NUMBER OF SEQ ID NOS: 285684  
SEQ ID NO 199559  
LENGTH: 524  
TYPE: PRT  
ORGANISM: Glycine max  
FEATURE:  
OTHER INFORMATION: Clone ID: PAT\_MRT3847\_22226C.1.pcp  
US-10-424-599-199559

Query Match 41.2%; Score 1124; DB 12; Length 524;  
 Best Local Similarity 40.5%; Pred. No. 5e-104;  
 Matches 213; Conservative 118; Mismatches 169; Indels 26; Gaps 6;

QY 8 WFIKPVLSVILSLVY---KMSLLMWRPKIEHHSKQIGRPPIHFPIGVKELY 63  
 DB 5 WATSSSIVFVIVIALATSMAMRMNLMTIRPRERLLRBOGLQNPFRILLVGDKEIV 64  
 QY 64 GMLKASSHPMPFSNNILPRVLSFYHMRKIYATFLWFGPFSTLVADPDLIREIFSK 123  
 DB 65 KLQWEAREKPMNLHDLVPRFAHLHQSVALKHGNSPIWGPRTLLDPLIDVLYNK 124  
 QY 124 SEPRKNAHLVYKQLEDGLSLKGEKNAHNRKIISPTFMENLKLVPVYLSKVTMV 183  
 DB 125 ISDFKRPANPLAKLL-ATGLVNYDGKMKHRRLLINPAFSLEKLMPIFFKSCNDLI 183  
 QY 184 DKMSDKLSENGEVEVDYEMFQILTEVDISRTAFSGSYEDGRAVFLQAQOMLCAEAFQ 243  
 DB 184 IKMGMLSYDSCENDVWPFQNLASDVIRTAFGSSFBEGKRIPOLOKELALMTXIM 243  
 QY 244 KYFLPGYRPPPTRGNLKSRKIDKEIRKSLKLIERRONALDGESECKEPAKOLLGLM 303  
 DB 244 KYFLPGWRFVPTATRRMKKEIDRYIKASLTDMIKKREKAPKTGEA-----TRDILLGL 297  
 QY 304 IOAKNVTVQ-----DIVECKSFPPAGKQTSNLLTWTILLSMHPWAKA 350  
 DB 298 LESNKEIOEHRNNENGMANLDYBECKLPTAGCETTSVLLWTVMLLSKRPMSRA 357  
 QY 351 RDEVLRVCGSHDVPKDHVVKLTLSMI-INESLRYPPIVATIRBAKDVYKGGYKIPC 409  
 DB 358 REEVLYQVFG-KQANFPGLSHLKIVTWIFLKVLRYPVAVGLNRRVDRMDKLGMLSDPA 416  
 QY 410 GTELLIPILAAHNOAIGNDVNEFNBARFADVPRAAKHPVGFIPGLGVRTCIQONILA 469  
 DB 417 GYQVSLPTTMVPHRELMDGVNEFFKERSSEGLATKNGRSFFPFPGPRICIGNFS 476  
 QY 470 ILQAKLTAAVMIQRTFHLAPYQAHPTVLMLLYPQHGAPIFRRL 515  
 DB 477 LLEAKMALSTILQHSFSELSPAVYAHAVTVFTLQPYGAHVLRFKY 522

RESULT 9  
 US-10-424-599-269344  
 ; Sequence 269344, Application US/104245599  
 ; Publication No. US20040031072A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: La Rosa Thomas J  
 ; APPLICANT: Kovalic David K  
 ; APPLICANT: Zhou Yihua  
 ; APPLICANT: Cao Yongwei  
 ; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
 ; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
 ; FILE REFERENCE: 38-21(53223)B  
 ; CURRENT APPLICATION NUMBER: US/10/424,599  
 ; CURRENT FILING DATE: 2003-04-28  
 ; NUMBER OF SEQ ID NOS: 285684  
 ; SEQ ID NO 269344  
 ; LENGTH: 533  
 ; TYPE: PRT  
 ; ORGANISM: Glycine max  
 ; FEATURE:  
 ; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_85238C.1.pep  
 US-10-424-599-269344

Query Match 40.8%; Score 1114; DB 12; Length 533;  
 Best Local Similarity 40.8%; Pred. No. 5.3e-103;  
 Matches 210; Conservative 118; Mismatches 167; Indels 20; Gaps 6;

QY 13 VLIVSVILSVIKGMSLLMWRPKIEHHSKQIGRPPIHFPIGVKELYGMLKASSH 72  
 DB 27 ITVATVLIWFMNALMVMVLPRIKRIERLKEOIGQNSYRPLIGDIRMDVAKIEAKSK 86  
 QY 73 PM-PRSHNILPRVLSFYHMRKIYGAFLVWFGPFTRLTVADPDLIREIFSKSEFEYKNE 131

DB 87 PNDPNSNDLPRAVLPVYVHTLANKGSKSPMWLGSTPRVFLIDPPKFEKMAKTXDYDQKPD 146  
 QY 132 AHPYKQLEGDGLSLKGEKNAHNRKIISPTFMENLKLVPVYLSKVTMDVMDKDL- 190  
 DB 147 TSPFLKLL-ASGFAYDDDKMAKRIKIVSPAFNERKCLVPIPCQSCDILISKMSLLS 205  
 QY 191 SENGVEVDYEMFQILTEVDISRTAFSGSYEDGRAVFLQAQOMLCAEAFQKVPPIGY 250  
 DB 206 SNGSCDLVWPFQNVSDVLAAGFSSYQESKIFELQREMIQLTWTLFFFAFIPGY 265  
 QY 251 RPPPTRGNLKSRKIDKEIRKSLKLIERRONALDGESECKEPAADILGMLTQAK--- 307  
 DB 266 RFLPHTMRKMAIDKEIRESLMVIINERLKAIRAG-----EFTNDLLGILLESYKE 319  
 QY 308 -----NVTVDIVECKSFPPAGKQTSNLLTWTILLSMHPWAKADEVLRYCS 360  
 DB 320 SEKSGGGSLSLEVEEYKLYLQGEANMELVWTLILLSHDPDQEKAGEVQVFGN 379  
 QY 361 RDPVTKDHVVKLTLSMIINESLRYPPIVATIRBAKDVYKGGYKIPCGTELLIPILAY 420  
 DB 380 -EKPDYERIGQIKIVSMILQESLRYPPVMAFAYLRKDKLTLPAGVELVAPVSKL 438  
 QY 421 HHDQAIWGNVDVNEFNBARFADVPRAAKHPVGFIPGLGVRTCIQONILALQAKTLAVN 480  
 DB 439 HODXEFMGDDAGEFPRFSEGVSKATKGLSYLPFGWGRRLCIGNFGLLAKVAVSMI 498  
 QY 481 IQRTFHLAPYQAHPTVLMLLYPQHGAPIFRRL 515  
 DB 499 LQRFSLHSPSYAHAPSPFITLQPERGAHILRL 533

RESULT 10  
 US-10-425-114-38180  
 ; Sequence 38180, Application US/10425114  
 ; Publication No. US2004003488BA1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Liu, Jingdong  
 ; APPLICANT: Zhou, Yihua  
 ; APPLICANT: Kovalic, David K.  
 ; APPLICANT: Screen, Steven E  
 ; APPLICANT: Tabaska, Jack E  
 ; APPLICANT: Cao Yongwei  
 ; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
 ; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
 ; FILE REFERENCE: 38-21(53313)B  
 ; CURRENT APPLICATION NUMBER: US/10/425,114  
 ; CURRENT FILING DATE: 2003-04-28  
 ; NUMBER OF SEQ ID NOS: 73128  
 ; SEQ ID NO 38180  
 ; LENGTH: 537  
 ; TYPE: PRT  
 ; ORGANISM: Zea mays  
 ; FEATURE:  
 ; OTHER INFORMATION: Clone ID: LIB3608-017-E5\_FIL.pep  
 US-10-425-114-38180

Query Match 40.4%; Score 1103.5; DB 12; Length 537;  
 Best Local Similarity 41.7%; Pred. No. 6.1e-102;  
 Matches 217; Conservative 100; Mismatches 180; Indels 23; Gaps 4;

QY 13 VLIVSVILSVIKGMSL---LWWRPKIEHHSKQIGRPPIHFPIGVKELYGMLKAA 69  
 DB 24 LIYGVGLGALLMKARLILDLRMTBPRLEALAQGLRGTSYRFLTGDJREFYRSKERA 83  
 QY 70 SHHPMPF-SHNILPRVLSFYHMRKIYGAFLVWFGPFTRLTVADPDLIREISKSEFEY 128  
 DB 84 MARPLDLRCHDLAGVEVFIFGAVLDEKTCFTFGVPVATVADPDLADVANKFGHF 143  
 QY 129 KNEAHPYKQLEGDGLSLKGEKNAHNRKIISPTFMENLKLVPVYLSKVTMDVMDKSD 188  
 DB 144 EKPKFPALTKLFSDDVANHEKVKWKRRILINPAFHLERKLMIPASACSEELVSRMAQ 203

```

QY 189 KLSNGEVDVYEMFQILTEDVISRTAFSSSYEDGRAVFRLOAQOMLLCAEAFQKVP 248
DB 204 SLGDSGCEHLDVDELQTLTGVDVTSRTAFSSSYLEGRKIFQCAEQERLMSIIDKFAVP 263
QY 249 GYRFPPTRGNLKSKRLDKETIRKSLKTIERRONALDGBECKEPAADLLGLMIOA-- 306
DB 264 GYMSLPTKNNRRMQIASEIDSLIRGLIGRQMAKQGBSDK-----DDLGLLHESNA 317
QY 307 -----KNTVODIVECKSPFFACKQTTSNLLTWTTLLSMHEMOWAKADEYL 355
DB 318 RETGDSGQPGQGLTMEBWECKLFFPAGMETTSVLLTWTTLLSMHEMOWDRAREYL 377
QY 356 RVCGSRDVPYKDHVVKLTSLMINSRLPYPIVATIRAKSDVLYGKIPCGTELLI 415
DB 378 GLFGKQPGYDGLSRKTYTMTLYEVLRLYPALAFSRKTYKEMVGVTPAGVTLEL 437
QY 416 PLIVHDDQAIWGNVNEFNPARFADVPRAAKHPVGFIPGLGVRTICIGNLAILOAKL 475
DB 438 PVLFIHDDPDIWSDAHEFRERAEVAAASDKRLAFFPFGMPRLCIGONFALLEAKM 497
QY 476 TLAVMIQRTFHLAPYQHAFTVMLLYPOHGAPITFRRL 515
DB 498 ALSMILQRFQELAPTYHVRVYIMLRPMHGAQIKLRAI 537

RESULT 11
US-10-424-599-217648
; Sequence 217648, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J
; APPLICANT: Kovalic, David K
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 217648
; LENGTH: 524
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_38563C.1.pep
US-10-424-599-217648

Query Match 40.4%; Score 1102; DB 12; Length 524;
Best Local Similarity 41.5%; Pred. No. 8.3e-102;
Matches 216; Conservative 112; Mismatches 165; Indels 28; Gaps 7;

QY 14 LVLSVILSVLYKGSLLMWRPRKIEEHFSGKGRGPPYHFFIGNVKELYGMMLKASSHP 73
DB 13 LIIAPLVLMCKLKWMLRPRKLERLRKGLGNFYSLLIDTEMVTVLMQARSQ 72
QY 74 MPEFS-----HNILPRVLSFYHWRKIYGATFLVWGFPTRLTVADPDILREIFSKSEFY 127
DB 73 QSTSFSLSKDOAAHITTFNHNHIVNFKGNSFFWEGTQPKVITIDPQILEVFNKIQDF 132
QY 128 EKNEAHPLYKQLEGDGLSLKGEKAAHRRKTIISPTFHEMNLKLVPLVYKSVPMVDKMS 187
DB 133 EKPKSPIVKLL-GSGLANLSEGEKWRTHKTIINPFLHLEKLVMLPITLSCCDMVKME 191
QY 188 KLSNGEVDVYEMFQILTEDVISRTAFSSSYEDGRAVFRLOAQOMLLCAEAFQKVP 247
DB 192 RLSSNDKSEIDVWPLQNLCDIISRTAFSSSYEDGRIFFELKEQGLMMK-LQNAVYI 250
QY 248 PGYRFPPTRGNLKSKRLDKETIRKSLKTIERRONALDGBECKEPAADLLGLMIOAK 307
DB 251 PGWMLPTTNNKMKKIDTEIR-ALKGVINREYNA-----KAGEVLNDDLGLMLBSN 304
QY 308 NVTVOD-----IVECKSPFFAGKQTTSNLLTWTTLLSMHEMOWAKADEV 354

```

```

DB 305 RMEIQHGNKNTIAMSLSVIECAFIYAGGETTSVLLTWTTLLSMHEMOWAKADEV 364
QY 355 LRVCGSRDVPYKDHVVKLTSLMINSRLPYPIVATIRAKSDVLYGKIPCGTELLI 414
DB 365 LHVFGNQK-PDYVGLSHLKIYTMILYEVRLPYPIVAFARLKNVVKGNLSLPGVQVS 423
QY 415 IPIIVHDDQAIWGNVNEFNPARFADVPRAAKHPVGFIPGLGVRTICIGNLAILOAK 474
DB 424 LPIILLHQDDIDWDATERPBERPABGAKATKQGVFFPGMKPRVCLGONFALLEAK 483
QY 475 TLAVMIQRTFHLAPYQHAFTVMLLYPOHGAPITFRRL 515
DB 484 LVLSLQRFSELSPTVAHAPVTYVTLNPKFGAHIILHKL 524

RESULT 12
US-10-424-599-228819
; Sequence 228819, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J
; APPLICANT: Kovalic, David K
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 228819
; LENGTH: 520
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_48650C.1.pep
US-10-424-599-228819

Query Match 37.5%; Score 1024; DB 12; Length 520;
Best Local Similarity 39.4%; Pred. No. 6.1e-94;
Matches 207; Conservative 109; Mismatches 185; Indels 24; Gaps 6;

QY 5 SSSWFIKVLVLSVILSVLYKGSLLMWRPRKIEEHFSGKGRGPPYHFFIGNVKELYG 64
DB 6 SSSAALCVTVLVAVITVLMKLNLSIMLRPKFELRLRQGHGPDYLSHHNSQTL 65
QY 65 MMLKASSHPMPSHNLPRVLSFYHWRKIYGATFLVWGFPTRLTVADPDILREIFSKS 124
DB 66 QQHQLNSQPPPLSDVDVAPRVSSLLHRTIDKYGKSSFFWEGRTPKVITIDPQILEVFN 125
QY 125 EYENKNAHPLYKQLEGDGLSLKGEKAAHRRKTIISPTFHEMNLKLVPLVYKSVPMVD 184
DB 126 HDQPKRSDNVKFLFA-GLINWGDKMAHRRKTMNPAFSEKLMPLRFSQSCDMIS 184
QY 185 KMSDLSNGEVDVYEMFQILTEDVISRTAFSSSYEDGRAVFR-LQAOQMLCAEAFQ 243
DB 186 MKKGLSSDGCHEIDIMFQNLTRDVISOAFSSSYAEKEKFRNLMQGYLLMAGKYK 244
QY 244 KVFIPGYRFPPTRGNLKSKRLDKETIRKSLKTIERRONALDGBECKEPAADLLGLM 303
DB 245 N-IPILNHLRTTTTKRMEALERIRISIEGIIKREKAWENG-----ETSNEDLISIL 296
QY 304 ID-----AKNVTVODIVECKSPFFACKQTTSNLLTWTTLLSMHEMOWAKA 350
DB 297 LESNHEKEQSGNSGAIVMTQGVIEBEKLYLAGQSTSSLLVWTTVLLAARYBEMOARA 356
QY 351 RDEVLRVCGSRDVPYKDHVVKLTSLMINSRLPYPIVATIRAKSDVLYGKIPCG 410
DB 357 RDEVQVFGNGON-PNIDISRLKIVTMILYEVLPPTTFPSAPQKDYKGLSLPAG 415
QY 411 TELLPILIVHDDQAIWGNVNEFNPARFADVPRAAKHPVGFIPGLGVRTICIGNLAI 470

```

```

Db      416 IRTMPLFIHDDGIDWDGDAKEFKERESGIIAKATKQISFYFPGMGPRICIGONRAL 475
Qy      471 LQAKTLAVMIQRFTHLAPTYOHAFTVLMILYPOHAGDITERRL 515
      476 MEAKTVLSILLQHFSELSFVYEHADTVLSIQPKGAHIVHKL 520

RESULT 13
US-10-425-114-50377
; Sequence 50377, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jindong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(5313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 50377
; LENGTH: 458
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3060-044-A6_FLI.pep
US-10-425-114-50377

Query Match      36.2%; Score 989; DB 12; Length 458;
Best Local Similarity 41.2%; Pred. No. 1.7e-90;
Matches 191; Conservative 92; Mismatches 159; Indels 22; Gaps 6;

Qy      67 LKASHPMP-----FSHNILPRVLSFYHMRKIYGATFLVWFGPTFLVADPDLIR 118
Db      2  LLALHPRPPPCGPRRLASTAPRVTPFMRRLVLEHGRTSMGSPSKTIIVPDLAK 61
Qy      119 EIRF-KSEFYKNAHFLVQLBEDGLSLKGEKMAHRIKISPTFMENLKLVPVYLK 177
Db      62 DVLNKGHFPEKLVPLSKVL-GSGVASHGEKVMYHRRILHAFHLEKLMKLPAST 120
Qy      178 SVTMVMDKM-SDKLSENGEVEVDYEWFOILTEDVISRTAFSGSYEDGSAVRLQAQML 236
Db      121 CCEELVSRMAESLGSOSCELDVMPRLQNLTDIISRITVFSYSYSEGRIRIQLQVEQAS 180
Qy      237 LCAEAFQKVFIPGRFPPTKGNLKSRLDYEIRKSLIKLIERRQNALIDGEGECKEPAA 296
Db      181 LMTNIRKIMIPGYMYLPTANRKMNRNREVSILREIIGKRIQAMEQGEGRK----- 234
Qy      297 KDILGIMTQAK-----NVTVDIYECKSPFPAGKOTTSNLLTWITILSMHEMQAKAR 351
Db      235 DDLVGLSLSTNRDMDMGWITIEDVIECKVFYFGMETTVLLMTWTVLSMHEMQDRAR 294
Qy      352 DEVLAVCGSRDVPFKDHYVVKLTLSMILNESLRLYPPIVATIRPAKSDVVLGGYKIPCGT 411
Db      295 ECVTLAFGRDCKPEYDGLSKLVTTMYLVEVLRILYRPATSVVROQYKEMEVGVTVYAGV 354
Qy      412 ELLPIIIVHDDQAIWGDVNEFNPAFADGVPRAAHGVGFIIPFGIGVTCIGONLALIQ 471
Db      365 ILLEPVLIHDDPDIWGDARFPRDRPSGVSASASCDPAFIIPFGISPRICIGONFALL 414
Qy      472 QAKTLAVMIQRFTHLAPTYOHAFTVLMILYPOHAGDITERRL 515
Db      415 EAKVALCMILQRFEGGLAPSYAHAPHMTITLHPHGAQLKRLAI 458

RESULT 14
US-10-425-114-58587
; Sequence 58587, Application US/10425114
; Publication No. US20040034888A1

```

```

; GENERAL INFORMATION:
; APPLICANT: Liu, Jindong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(5313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 58587
; LENGTH: 532
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: 700076168_FLI.pep
US-10-425-114-58587

Query Match      35.7%; Score 973.5; DB 12; Length 532;
Best Local Similarity 39.0%; Pred. No. 7.8e-89;
Matches 204; Conservative 107; Mismatches 181; Indels 31; Gaps 9;

Qy      15 VLSTVILSVLVK-GMSIL---WVRPKLIEHFSKQIGRPYHPFPGIVKELVGMMLKAS 70
Db      19 LLYVLSGLVALRLAWQVLEWGLSPRLRLRALRAEGLRGTAVFPAGAKBERLLAAAR 78
Qy      71 SHMP-FSNHILPRVLSFYHMRKIYGATFLVWFGPTFLVADPDLIRFISKS-----E 125
Db      79 AKPMPISHTAISARVAPLVNNAIHERKISVMIGITPKVILSDKLVREVSKSKSGCD 138
Qy      126 FYENEAHPLVQLBEDGLSLKGEKMAHRIKISPTFMENLKLVPVYLKSVTMDYDK 185
Db      139 FQPKPLSHRIK-LIQGLVTHEGECWALHRIKISAFVYELKEMLPATFACGTGELMGR 197
Qy      186 WSDKLSNGEVEVDYEWFOILTEDVISTAGSSYEDGSAVRLQAQMLCAEAFQV 245
Db      198 WESMGDDGAREIDVWPEQDITGDALISPAAGSSILIBERRILFQAQIKI-ANEMTVL 256
Qy      246 FIDGYRFPPTKGNLKSRLDYEIRKSLIKLIERRQNALIDGEGECKEPAAIDLGLMIQ 305
Db      257 YIEGYTLTPFKRRRIKARAEVALLRGIITKRETRMNGVAD-----SDLIGKLMQ 310
Qy      306 AK-----NVTVDIYECKSPFPAGKOTTSNLLTWITILSMHEMQAKAR 352
Db      311 SNVEAHGGGSSKPMWTMEDILGKLFYFGMETTSALLMTWITVLSMHEMQDRAR 370
Qy      353 EVLAVCGSRDVPFKDHYVVKLTLSMILNESLRLYPPIVATIRPAKSDVVLGGYKIPCGT 412
Db      371 EVLAVFGKME-PVINGNSLKVTTMYLVEVLRILYRPIDLEQYKEMELGGVARYPGVT 429
Qy      413 LLIPIIIVHDDQAIWGDVNEFNPAFADGVPRAAHGVGFIIPFGIGVTCIGONLALIQ 472
Db      430 LLPIIILIHDDPDIWGDVQFRRPFRFADGISASASDTPAFRPFPGIPRIQVQNFALLE 489
Qy      473 AKTLAVMIQRFTHLAPTYOHAFTVLMILYPOHAGDITERRL 515
Db      490 AKVALAMLLQRFSGLSPTTHAPFSVSTVQPSHQAQIVYKXI 532

RESULT 15
US-10-425-114-61220
; Sequence 61220, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jindong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei

```



TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
FILE REFERENCE: 38-21(5313)B  
CURRENT APPLICATION NUMBER: US/10/425,114  
CURRENT FILING DATE: 2003-04-28  
NUMBER OF SEQ ID NOS: 73128  
SEQ ID NO 61220  
LENGTH: 540  
TYPE: PRT  
ORGANISM: Zea mays  
FEATURE:  
OTHER INFORMATION: Clone ID: LIB143-011-G1\_F11.pap  
US-10-425-114-61220

Query Match 35.4%; Score 967; DB 12; Length 540;  
Best Local Similarity 37.7%; Pred. No. 3.6e-88;  
Matches 205; Conservative 106; Mismatches 181; Indels 52; Gaps 9;

QY 7 SWPIPKVLVSLVSLVYKMSLM-----WRPKIEHFSKQIGRPYHPTGN 58  
DB 13 SMAFPAYPLAALLVATR---LWDYFVRLVWRPVAITKGFEQIGHSYRFFKGC 68  
QY 59 VKELVGMMLKASHPMF-SHNLPRVLSYHHMKIGATFLVWGPFFRLTVADPLI 117  
DB 69 NEELRSMKERTDGLVDVGHKYLPRAPYLEMRAQYGPFLYWGQARICIDYELA 128  
QY 118 REIF-SKSEFEYKKEAHPVLEKQEGDLSLKGEKVAHHRKIIISPTFHENLKLVPVL 176  
DB 129 KQISSKSGHVRVKNDAFTLLALVGKLGFMESDWRHRVRVNPFTIDKLK---IVT 184  
QY 177 KSVTDMVDKMSDKL-----SENGEVVDVYEMFQILTEDIISRTAFSSYEDGRVFL 230  
DB 185 ETMLDPADSMAGELIAPASQENGETQVDYKHSDDLVTNLAVALFGSSYKLGKQVEA 244  
QY 231 QAQMMLCAEAFQVFIPIGYRFPFTGNLKSRLDKERKSLKLIERRRONAIDEGEE 290  
DB 245 QTELGITMAFTLDVPIPGSKYLPLOANRKNMLETKXSLTRIIQPRLASGEGH--- 300  
QY 291 CKEPANDLGLMTQA-----KNTVQDIVECKSPFPAKQITSNLL 333  
DB 301 -----NDLGVMLDSCTETKQGGKQOQVDRRLSLSMEEIIECKLFFAGHENTALL 354  
QY 334 TWTITLSMHPWQAKARDEVLRVCGSRDVPYKDVVVKLTLSMILNESLRLYPIVATI 393  
DB 355 TWSVYLSTIPEWQERLKEVLRFEFG-RDAPNPAALNRLEKMTMVLFTLRILSPALFMQ 413  
QY 394 RRAKSDVLAGYKIPCGTELLIPPIAVHDAQIWNQDVNEFNFPADGVPRAKHPVGF 453  
DB 414 RKTADMTVGPIKLPKGTALVPIPIVHRDXQAWGDADDEFSPRFANGITGAAPVHGL 473  
QY 454 IPIGIVATCTGGUATLAIQALTLAVMIQRTFHLAPTQHAPTVLMLYPOHAPITFR 513  
DB 474 LAFSGPRSCIGNLSMLEAKSTLALMKRKFSPALSPDYVAHPVDLFTLKPKGLPVILR 533  
QY 514 RLIN 517  
DB 534 PLDN 537

Search completed: May 19, 2004, 10:33:19  
Job time : 50 secs